(12) NACH DEM VERTRAG ÜBER DIE INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES PATENTWESENS (PCT) VERÖFFENTLICHTE INTERNATIONALE ANMELDUNG

(19) Weltorganisation für geistiges Eigentum
Internationales Büro

(43) Internationales Veröffentlichungsdatum
1. April 2004 (01.04.2004)

PCT

(10) Internationale Veröffentlichungsnummer
WO 2004/027069 A1

(51) Internationale Patentklassifikation⁷: C12N 15/82, 15/11, 15/63, A01K 67/027

(21) Internationales Aktenzeichen: PCT/EP2003/008394

(22) Internationales Anmeldedatum: 30. Juli 2003 (30.07.2003)

(25) Einreichungssprache: Deutsch

(26) Veröffentlichungssprache: Deutsch

(30) Angaben zur Priorität:
102 38 979.9 20. August 2002 (20.08.2002) DE
102 47 599.7 11. Oktober 2002 (11.10.2002) DE

(71) Anmelder (*für alle Bestimmungsstaaten mit Ausnahme von US*): SUNGENE GMBH & CO. KGAA [DE/DE]; Corrensstr. 3, 06466 Gatersleben (DE).

(72) Erfinder; und

(75) Erfinder/Anmelder (*nur für US*): SCHOPFER, Christel, Renate [DE/DE]; Konvent 38, 06484 Quedlinburg (DE). SAUER, Matt [DE/DE]; Markt 9, 06484 Quedlinburg (DE). KLEBSATTEL, Martin [DE/DE]; Weingarten 9, 06484 Quedlinburg (DE). FLACHMANN, Ralf [DE/DE]; Halberstädter Str. 20a, 06484 Quedlinburg (DE).

(74) Anwalt: DÖRPER, Thomas; c/o BASF Aktiengesellschaft, 67056 Ludwigshafen (DE).

(81) Bestimmungsstaaten (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Bestimmungsstaaten (*regional*): ARIPO-Patent (BW, GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), eurasisches Patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), europäisches Patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI-Patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Veröffentlicht:
— *mit internationalem Recherchenbericht*

Zur Erklärung der Zweibuchstaben-Codes und der anderen Abkürzungen wird auf die Erklärungen ("Guidance Notes on Codes and Abbreviations") am Anfang jeder regulären Ausgabe der PCT-Gazette verwiesen.

WO 2004/027069 A1

(54) Title: TRANSGENIC EXPRESSION CASSETTES FOR THE EXPRESSION OF NUCLEIC ACIDS IN PLANT BLOOMS

(54) Bezeichnung: TRANSGENE EXPRESSIONSKASSETTEN ZUR EXPRESSION VON NUKLEINSÄUREN IN DER PFLANZLICHEN BLÜTE

(57) Abstract: The invention relates to methods for the targeted, transgenic expression of nucleic acid sequences in plant blooms and transgenic expression cassettes and expression vectors which have promoters with an expression specificity for plant blooms. The invention further relates to organisms modified with said transgenic expression cassettes or expression vectors (preferably plants), cultures derived therefrom, parts or propagation material and the use of the above for the production of human and animal foodstuffs, seed stock, pharmaceuticals or fine chemicals.

(57) Zusammenfassung: Die Erfindung betrifft Verfahren zur gezielten, transgenen Expression von Nukleinsäuresequenzen in der pflanzlichen Blüte, sowie transgene Expressionskassetten und Expressionsvektoren, die Promotoren mit einer Expressionsspezifität für die pflanzliche Blüte enthalten. Die Erfindung betrifft ferner mit diesen transgenen Expressionskassetten oder Expressionsvektoren transformierte Organismen (bevorzugt Pflanzen), davon abgeleitete Kulturen, Teile oder Vermehrungsgut, sowie die Verwendung derselben zur Herstellung von Nahrungs-, Futtermitteln, Saatgut, Pharmazeutika oder Feinchemikalien.

(19) Canadian Intellectual Property Office
An Agency of Industry Canada

Office de la Propri,t, Intellectuelle du Canada
Un organisme d'Industrie Canada

(11) CA 2 496 300 (13) A1

(40) 01.04.2004
(43) 01.04.2004

(12)

(21) 2 496 300

(22) 30.07.2003

(51) Int. Cl.7: C12N 15/82, A01K 67/027, C12N 15/11, C12N 15/63

(85) 18.02.2005

(86) PCT/EP03/008394

(87) WO04/027069

(30) 102 38 979.9 DE 20.08.2002
102 47 599.7 DE 11.10.2002

(71) SUNGENE GMBH & CO. KGAA,
Correnssstr. 3, 06466, GATERSLEBEN, XX (DE).

(72) SAUER, MATT (DE).
SCHOPFER, CHRISTEL RENATE (DE).
FLACHMANN, RALF (DE).
KLEBSATTEL, MARTIN (DE).

(74) ROBIC

(54) CASSETTES D'EXPRESSION TRANSGENIQUE POUR L'EXPRESSION D'ACIDES NUCLEIQUES DANS UNE FLEUR VEGETALE

(54) TRANSGENIC EXPRESSION CASSETTES FOR THE EXPRESSION OF NUCLEIC ACIDS IN PLANT BLOOMS

(57)

The invention relates to methods for the targeted, transgenic expression of nucleic acid sequences in plant blooms and transgenic expression cassettes and expression vectors which have promoters with an expression specificity for plant blooms. The invention further relates to organisms modified with said transgenic expression cassettes or expression vectors (preferably plants), cultures derived therefrom, parts or propagation material and the use of the above for the production of human and animal feedstuffs, seed stock, pharmaceuticals or fine chemicals.

Office de la Propriété Intellectuelle du Canada
Un organisme d'Industrie Canada

Canadian Intellectual Property Office
An agency of Industry Canada

CA 2496300 A1 2004/04/01

(21) 2 496 300

(12) DEMANDE DE BREVET CANADIEN
CANADIAN PATENT APPLICATION

(13) A1

(86) Date de dépôt PCT/PCT Filing Date: 2003/07/30
(87) Date publication PCT/PCT Publication Date: 2004/04/01
(85) Entrée phase nationale/National Entry: 2005/02/18
(86) N° demande PCT/PCT Application No.: EP 2003/008394
(87) N° publication PCT/PCT Publication No.: 2004/027069
(30) Priorités/Priorities: 2002/08/20 (102 38 979.9) DE;
2002/10/11 (102 47 599.7) DE

(51) Cl.Int.^{7}/Int.Cl.^{7} C12N 15/82, A01K 67/027, C12N 15/63, C12N 15/11

(71) Demandeur/Applicant:
SUNGENE GMBH & CO. KGAA, DE

(72) Inventeurs/Inventors:
SCHOPFER, CHRISTEL RENATE, DE;
SAUER, MATT, DE;
KLEBSATTEL, MARTIN, DE;
FLACHMANN, RALF, DE

(74) Agent: ROBIC

(54) Titre : CASSETTES D'EXPRESSION TRANSGENIQUE POUR L'EXPRESSION D'ACIDES NUCLEIQUES DANS UNE FLEUR VEGETALE

(54) Title: TRANSGENIC EXPRESSION CASSETTES FOR THE EXPRESSION OF NUCLEIC ACIDS IN PLANT BLOOMS

(57) Abrégé/Abstract:
The invention relates to methods for the targeted, transgenic expression of nucleic acid sequences in plant blooms and transgenic expression cassettes and expression vectors which have promoters with an expression specificity for plant blooms. The invention further relates to organisms modified with said transgenic expression cassettes or expression vectors (preferably plants), cultures derived therefrom, parts or propagation material and the use of the above for the production of human and animal feedstuffs, seed stock, pharmaceuticals or fine chemicals.

(12) NACH DEM VERTRAG ÜBER DIE INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES PATENTWESENS (PCT) VERÖFFENTLICHTE INTERNATIONALE ANMELDUNG

(19) Weltorganisation für geistiges Eigentum
Internationales Büro

(43) Internationales Veröffentlichungsdatum
1. April 2004 (01.04.2004)

PCT

(10) Internationale Veröffentlichungsnummer
WO 2004/027069 A1

(51) Internationale Patentklassifikation[7]: C12N 15/82, 15/11, 15/63, A01K 67/027

(21) Internationales Aktenzeichen: PCT/EP2003/008394

(22) Internationales Anmeldedatum:
30. Juli 2003 (30.07.2003)

(25) Einreichungssprache: Deutsch

(26) Veröffentlichungssprache: Deutsch

(30) Angaben zur Priorität:
102 38 979.9 20. August 2002 (20.08.2002) DE
102 47 599.7 11. Oktober 2002 (11.10.2002) DE

(71) Anmelder *(für alle Bestimmungsstaaten mit Ausnahme von US)*: **SUNGENE GMBH & CO. KGAA** [DE/DE]; Corrensstr. 3, 06466 Gatersleben (DE).

(72) Erfinder; und
(75) Erfinder/Anmelder *(nur für US)*: **SCHOPFER, Christel, Renate** [DE/DE]; Konvent 38, 06484 Quedlinburg (DE). **SAUER, Matt** [DE/DE]; Markt 9, 06484 Quedlinburg (DE). **KLEBSATTEL, Martin** [DE/DE]; Weingarten 9, 06484 Quedlinburg (DE). **FLACHMANN, Ralf** [DE/DE]; Halberstädter Str. 20a, 06484 Quedlinburg (DE).

(74) Anwalt: **DÖRPER, Thomas**; c/o BASF Aktiengesellschaft, 67056 Ludwigshafen (DE).

(81) Bestimmungsstaaten *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Bestimmungsstaaten *(regional)*: ARIPO-Patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), eurasisches Patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), europäisches Patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI-Patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Veröffentlicht:
— *mit internationalem Recherchenbericht*

Zur Erklärung der Zweibuchstaben-Codes und der anderen Abkürzungen wird auf die Erklärungen ("Guidance Notes on Codes and Abbreviations") am Anfang jeder regulären Ausgabe der PCT-Gazette verwiesen.

(54) Title: TRANSGENIC EXPRESSION CASSETTES FOR THE EXPRESSION OF NUCLEIC ACIDS IN PLANT BLOOMS

(54) Bezeichnung: TRANSGENE EXPRESSIONSKASSETTEN ZUR EXPRESSION VON NUKLEINSÄUREN IN DER PFLANZLICHEN BLÜTE

(57) Abstract: The invention relates to methods for the targeted, transgenic expression of nucleic acid sequences in plant blooms and transgenic expression cassettes and expression vectors which have promoters with an expression specificity for plant blooms. The invention further relates to organisms modified with said transgenic expression cassettes or expression vectors (preferably plants), cultures derived therefrom, parts or propagation material and the use of the above for the production of human and animal feedstuffs, seed stock, pharmaceuticals or fine chemicals.

(57) Zusammenfassung: Die Erfindung betrifft Verfahren zur gezielten, transgenen Expression von Nukleinsäuresequenzen in der pflanzlichen Blüte, sowie transgene Expressionskassetten und Expressionsvektoren, die Promotoren mit einer Expressionsspezifität für die pflanzliche Blüte enthalten. Die Erfindung betrifft ferner mit diesen transgenen Expressionskassetten oder Expressionsvektoren transformierte Organismen (bevorzugt Pflanzen), davon abgeleitete Kulturen, Teile oder Vermehrungsgut, sowie die Verwendung derselben zur Herstellung von Nahrungs-, Futtermitteln, Saatgut, Pharmazeutika oder Feinchemikalien.

WO 2004/027069 A1

TRANSGENIC EXPRESSION CASSETTES FOR THE EXPRESSION OF NUCLEIC ACIDS IN PLANT BLOOMS

The invention relates to methods for the targeted transgenic expression of nucleic acid sequences in the flower of plants, and to transgenic expression cassettes and expression vectors which comprise promoters having an expression specificity for the flower of plants. The invention further relates to organisms (preferably plants) transformed with these transgenic expression cassettes or expression vectors, to cultures, parts or propagation material derived therefrom, and to the use of the same for producing human and animal foods, seeds, pharmaceuticals or fine chemicals.

The aim of biotechnological operations on plants is to produce plants with advantageous novel properties, for example for increasing the agricultural productivity, for increasing the quality of human foods or for producing particular chemicals or pharmaceuticals (Dunwell JM (2000) J Exp Bot 51 Spec No:487-96). A basic precondition for transgenic expression of particular genes is the provision of promoters which are functional in plants. Promoters are important tools in plant biotechnology for controlling the expression of particular genes in a transgenic plant and thus achieving particular traits of the plant.

Various promoters functional in plants are known, for example constitutive promoters such as the promoter of the agrobacterium nopaline synthase, the TR double promoter or the promoter of the cauliflower mosaic virus (CaMV) 35S transcript (Odell et al.(1985) Nature 313:810-812). A disadvantage of these promoters is that they are constitutively active in virtually all tissues of the plant. Targeted expression of genes in particular plant parts or at particular times of development is not possible with these promoters. There is thus a particularly great need for promoters having a defined activity profile and a specificity for particular plant tissues.

Promoters having specificities for various plant tissues such as anthers, ovaries, flowers, leaves, stalks, roots, tubers or seeds have been described. The stringency of the specificity and the expression activity of these promoters varies widely.

The flower of plants serves for sexual reproduction of flowering plants. The flowers of plants - especially the petals - frequently accumulate large amounts of secondary plant products such as, for example, terpenes, anthocyans, carotenoids,

alkaloids and phenylpropanoids, which serve as scents, defensive substances or as colorants in the flower. Many of these substances are of commercial interest. In addition, the flower bud and the flower of the plant is a sensitive organ, especially to stress factors such as cold.

Flower-specific promoters such as, for example, the phytoene synthase promoter (WO 92/16635), the promoter of the P-rr gene (WO 98/22593) or the promoter of the APETALA3 gene (Hill TA et al. (1998) Development 125:1711-1721) are known. However, all these promoters have one or more disadvantages which are predjudicial to wide use:

1) within the flower they are specific for one or more flower tissues and do not guarantee expression in all tissues of the flower.

2) they are - as in the example of the APETALA3 gene which is involved in flower development - highly regulated during flower development and are not active in all phases of flower development.

3) they occassionally show strong secondary activities in other plant tissues.

Despite the large number of known plant promoters, there is a need for promoters having a specificity for the flower of plants and guaranteeing high expression over a long period of flower development and flowering.

It is an object of the present invention to provide methods and suitable promoters for the targeted transgenic expression of nucleic acids in flower tissues.

We have found that this object is achieved by providing promoters of ε-cyclase. These promoters show an usually strong expression in numerous flower organs.

A first aspect of the invention relates to methods for the targeted transgenic expression of nucleic acid sequences in the flower of plants, including the following steps

I. introduction of a transgenic expression cassette into plant cells, where the transgenic expression cassette comprises at least the following elements

a) at least one promoter sequence of a gene coding for an ε-cyclase, and

b) at least one further nucleic acid sequence, and

c) where appropriate further genetic control elements,

where at least one of said promoter sequences and one further nucleic acid sequence are functionally linked together, and the further nucleic acid sequence is heterologous in relation to the promoter sequence or the plant cell, and

II. selection of transgenic cells which comprise said expression cassette stably integrated into the genome, and

III. regeneration of complete plants from said transgenic cells, where at least one of the further nucleic acid sequences is expressed in the flower.

A further aspect relates to transgenic expression cassettes as can be employed in the method of the invention. The transgenic expression cassettes preferably include for the targeted transgenic expression of nucleic acid sequences in the flower of plants

a) at least one promoter sequence of gene coding for an ε-cyclase, and

b) at least one further nucleic acid sequence, and

c) where appropriate further genetic control elements,

where at least one promoter sequence and one further nucleic acid sequence are functionally linked together, and the further nucleic acid sequence is heterologous in relation to the promoter sequence.

In a preferred embodiment of the method of the invention and/or of the expression cassettes of the invention, ''promoter sequence of a gene coding for an ε-cyclase'' means a sequence selected from the group of sequences consisting of

i) the promoter sequence of the ε-cyclase from Tagetes erecta as shown in SEQ ID NO: 1, the ε-cyclase from Arabidopsis thaliana as shown in SEQ ID NO: 7, the ε-cyclase from Oryza

sativa as shown in SEQ ID NO: 8, and

ii) functional equivalents of the promoter sequences as shown in SEQ ID NO: 1, 7 or 8 having substantially the same promoter activity as the promoter of the ε-cyclases as shown in SEQ ID NO: 1, 7 or 8 and

iii) functionally equivalent fragments of the sequences under i) or ii) having substantially the same promoter activity as the promoter of ε-cyclases as shown in SEQ ID NO: 1, 7 or 8.

It is particularly preferred for ''promoter sequence of a gene coding for an ε-cyclase'' to mean the promoter sequence from Tagetes erecta as shown in SEQ ID NO: 1 and functionally equivalent fragments thereof.

The expression cassettes of the invention may comprise further genetic control sequences and/or additional functional elements.

It is possible and preferred for the transgenic expression cassettes to make possible, through the nucleic acid sequence to be expressed transgenically, the expression of a protein encoded by said nucleic acid sequence and/or the expression of a sense-RNA, antisense-RNA or double-stranded RNA encoded by said nucleic acid sequence.

A further aspect of the invention relates to transgenic expression vectors which comprise one of the expression cassettes of the invention.

A further aspect of the invention relates to transgenic organisms which comprise one of the expression cassettes or expression vectors of the invention. The organism can be selected from the group consisting of bacteria, yeasts, fungi, nonhuman animals and plant organisms or of cells, cell cultures, parts, tissues, organs or propagation material derived therefrom, and the organism is preferably selected from the group of agricultural crop plants.

A further aspect of the invention therefore relates to an isolated nucleic acid sequence including the promoter of the ε-cyclase from Tagetes erecta as shown in SEQ ID NO: 1, and functionally equivalent fragments thereof.

In a preferred embodiment, the nucleic acid sequence of the invention or the transgenic expression cassette of the invention

in the form of a functionally equivalent promoter sequence includes, besides the sequence shown in SEQ ID NO: 1, additionally the sequence coding for the 5'-untranslated region of the ε-cyclase gene from Tagetes erecta. The sequence described by SEQ ID NO: 2 is particularly preferred.

In a further preferred embodiment, the nucleic acid sequence of the invention or the transgenic expression cassette of the invention in the form of a functionally equivalent promoter sequence includes, besides the sequence shown in SEQ ID NO: 1, additionally the sequence coding for the 5'-untranslated region of the ε-cyclase gene from Tagetes erecta and a sequence coding for a transit peptide, preferably for the transit peptide of the ε-cyclase protein from Tagetes erecta as shown in SEQ ID NO: 4. This sequence is preferably oriented in the 3' direction in relation to one of the promoters of the invention. The sequence described by SEQ ID NO: 3 is particularly preferred as promoter sequence in this connection.

A further aspect relates to the use of the isolated nucleic acid sequences, transgenic expression vectors or transgenic organisms of the invention for the transgenic expression of nucleic acids and/or proteins.

A further aspect of the invention relates to the use of the nucleic acid sequence of the invention for reducing the expression of an ε-cyclase. Included within this according to the invention are expression cassettes able to express a double-stranded RNA corresponding to the promoter sequence.

It is particularly preferred to use said transgenic organisms or cells, cell cultures, parts, tissues, organs or propagation material derived therefrom to produce human and animal foods, seeds, pharmaceuticals or fine chemicals, where the fine chemicals are preferably enzymes, vitamins, amino acids, sugars, saturated or unsaturated fatty acids, natural or synthetic flavorings, aromatizing substances or colorants. The invention further includes methods for producing said human and animal foods, seeds, pharmaceuticals or fine chemicals employing the trangenic organisms of the invention or cells, cell cultures, parts, tissues, organs or propagation material derived therefrom.

The transgenic expression cassettes of the invention are particularly advantageous for the following reason:

a) they impart selective expression in the flower of plant and make numerous applications possible, such as, for example, resistance to stress factors such as cold or targeted synthesis of secondary plant products. Expression takes place throughout the period of flower development with high activity.

''Expression'' means the transcription of the nucleic acid sequence which is to be expressed transgenically, but may also include - in the case of an open reading frame in the sense orientation - the translation of the transcribed RNA of the nucleic acid sequence to be expressed transgenically into a corresponding polypeptide.

''Transgenic'' means - for example in relation to a transgenic expression cassette, a transgenic expression vector, a transgenic organism or methods for the transgenic expression of nucleic acids - all constructions resulting from methods of genetic manipulation, or methods using such, in which either

a) an ε-cyclase promoter (e.g. as shown in SEQ ID NO: 1, 7 or 8) or a functional equivalent thereof or a functionally equivalent fragment of the aforementioned, or

b) the nucleic acid sequence which is to be transgenically expressed, are functionally linked to a promoter according to a), or

c) (a) and (b)

are not located in their natural genetic environment or have been modified by methods of genetic manipulation, where the modifications may be for example substitutions, additions, deletions, inversions or inserts of one or more nucleotide residues. The promoter sequence of the invention (e.g. the sequence as shown in SEQ ID NO: 1, 7 or 8) contained in the expression cassettes is preferably heterologous in relation to the further nucleic acid sequence which is to be expressed transgenically and is functionally linked thereto. ''Heterologous'' means in this connection that the further nucleic acid sequence does not code for the gene which is naturally under the control of said promoter.

''Natural genetic environment'' means the natural chromosomal locus in the original organism or the presence in a genomic library. In the case of a genomic library, the natural genetic

environment of the nucleic acid sequence is preferably still retained at least in part. The environment flanks the nucleic acid sequence at least on one side and has a sequence length of at least 50 bp, preferably at least 500 bp, particularly preferably at least 1000 bp, very particularly preferably at least 5000 bp. A naturally occurring expression constant - for example the naturally occurring combination of the promoter as shown in SEQ ID NO: 1 and of a gene coding for a protein as shown in SEQ ID NO: 10 or 12 - becomes a trangenic expression construct when the latter is modified by unnatural, synthetic (''artificial'') methods such as, for example, an in vitro mutagenesis. Appropriate methods are described (US 5,565,350; WO 00/15815; see also above).

''Transgenic'' means in relation to an expression (''transgenic expression'') preferably all expressions caused by use of a transgenic expression cassette, transgenic expression vector or transgenic organism - complying with the definitions given above.

The transgenic expression cassettes of the invention, and the transgenic expression vectors and transgenic organisms derived therefrom may include functional equivalents to the ε-cyclase promoter sequence described in SEQ ID NO: 1, 7 or 8.

Functional equivalents also include all the sequences derived from the complementary strand of the sequences defined by SEQ ID NO: 1, 7 or 8 and having substantially the same promoter activity. Particularly preferably included are the sequences shown in SEQ ID NO: 2 or 3, which, besides the promoter sequence, comprise the 5'-untranslated region or the 5'-untranslated region and the region coding for the transit peptide of the ε-cyclase from Tagetes erecta.

Functional equivalents means in particular natural or artificial mutations of the ε-cyclase promoter sequence described in SEQ ID NO: 1, 7 or 8, and the homologs thereof from other plant genera and species which still have substantially the same promoter activity as the ε-cyclase promoter as shown in SEQ ID NO: 1, 7 or 8.

A promoter activity is referred to as substantially the same when the transcription of a particular gene to be expressed is, under the control of, for example, a functional equivalent of the ε-cyclase promoter sequence described by SEQ ID NO: 1, 7 or 8, or of a functionally equivalent fragment thereof - under conditions

which are otherwise unchanged - higher in at least one flower tissue than in another non-flower tissue, for example the root or the leaves. In this connection, the expression under the control of one of the promoters of the invention in a flower tissue is preferably at least twice or five times, very particularly preferably at least ten times or fifty times, most preferably at least hundred times, that in another non-flower tissue, for example the root or the leaves.

''Flower'' generally means a shoot of limited growth whose leaves have been transformed into reproductive organs. The flower consists of various ''flower tissues'' such as, for example, the sepals, the petals, the stamens or the carpels. Androecium is the term used for the totality of stamens in the flower. The stamens are located within the circle of petals and sepals. A stamen is composed of a filament and of an anther located at the end. The latter in turn is divided into two thecae which are connected together by a connective. Each theca consists of two pollen sacs in which the pollen is formed.

''Targeted'' means in relation to expression in the flowers of plants preferably that the expression under the control of one of the promoters of the invention in at least one plant flower tissue is at least ten times, particularly preferably at least fifty times, very particularly preferably at least one hundred times that in a non-flower tissue such as, for example, the leaves.

The sequences preferably employed for estimating the level of expression are those which code for easily quantifiable proteins. Very particular preference is given in this connection to reporter proteins (Schenborn E, Groskreutz D. (1999) Mol Biotechnol 13(1): 29-44) such as the green fluorescent protein (GFP) (Chiu WL et al. (1996) Curr Biol 6:325-330; Leffel SM et al.(1997) Biotechniques 23(5):912-8), chloramphenicol acetyltransferase, luciferase (Millar et al. (1992) Plant Mol Biol Rep 10:324-414), ß-glucuronidase or β-galactosidase. Very particular preference is given to ß-glucuronidase (Jefferson et al. (1987) EMBO J 6:3901-3907).

''Conditions which are otherwise unchanged'' means that the expression initiated by one of the transgenic expression cassettes to be compared is not modified by combination with additional genetic control sequences, for example enhancer sequences. Unchanged conditions further means that all general conditions such as, for example, plant species, stage of

development of the plants, culture conditions, assay conditions (such as buffer, temperature, substrates etc.) are kept identical between the expressions to be compared.

Functional equivalents of the ε-cyclase promoter as shown in SEQ ID NO: 1, 7 or 8 preferably includes sequences which

a) have substantially the same promoter activity as the ε-cyclase promoter as shown in SEQ ID NO: 1, 7 or 8, and

b) have a homology of at least 50%, preferably 70%, more preferably at least 80%, particularly preferably at least 90%, very particularly preferably at least 95%, most preferably 99%, with the sequence of the ε-cyclase promoter shown in SEQ ID NO: 1, 7 or 8, where the homology extends over a length of at least 100 base pairs, preferably at least 200 base pairs, particularly preferably of at least 300 base pairs, very particularly preferably of at least 400 base pairs, most preferably of at least 500 base pairs.

It is possible in this connection for the level of expression of the functional equivalents to differ both downwards and upwards from a comparison value. Preference is given in this connection to the sequences whose level of expression, measured on the basis of the transcribed mRNA or the subsequently translated protein, under conditions which are otherwise unchanged differs quantitatively by not more than 50%, preferably 25%, particularly preferably 10%, from a comparison value obtained with the promoters described by SEQ ID NO: 1, 7 or 8. Particularly preferred sequences are those whose level of expression, measured on the basis of the transcribed mRNA or the subsequently translated protein, under conditions which are otherwise unchanged exceeds quantitatively by more than 50%, preferably 100%, particularly preferably 500%, very particularly preferably 1000%, a comparison value obtained with the promoter described by SEQ ID NO: 1, 7 or 8.

Further examples of functionally equivalent promoter sequences employed in the transgenic expression cassettes or transgenic expression vectors of the invention can easily be found for example in various organisms whose genomic sequence is at least partly known, such as, for example, Arabidopsis thaliana, Brassica napus, Nicotiana tabacum, Solanum tuberosum, Helianthus annuus, Linum sativum oder Oryza sativa, followed by homology comparisons in databases. A possible and preferred starting point for this is the coding regions of the gene whose promoter is

described by SEQ ID NO: 1, 7 or 8. Starting from, for example, the cDNA sequences of these genes described by SEQ ID NO: 9, 11, 13 or 15 or the protein sequence derived therefrom and described by SEQ ID NO: 10, 12, 14 or 16 it is possible easily to identify, in a manner familiar to the skilled worker, the corresponding homologous genes - and thus the relevant promoter regions, in other plant species by screening databases or gene libraries (using appropriate gene probes).

In a further preferred embodiment, functional equivalents to the ε-cyclase promoter as shown in SEQ ID NO: 1, 7 or 8 include sequences which are located in a plant organism in the 5' direction in front of a genomic sequence which codes for an ε-cyclase.

ε-Cyclase means in general all proteins which have an ε-cyclase activity.

By ε-cyclase activity is meant the enzymic activity of an ε-cyclase.

An ε-cyclase means a protein which has the enzymatic activity of converting a terminal linear lycopene residue into an ε-ionone ring.

In particular, ε-cyclase means in general all proteins able to catalyze the cyclization of lycopene to δ-carotene (and where appropriate further to ε-carotene) and/or of neurosporene to α-zeacarotene. The ε-cyclase preferably has an oxidoreductase activity and/or naturally shows a predominant localization in the plastids, especially the chloroplasts and chromoplasts.

An ε-cyclase preferably means a protein having the enzymatic activity for converting lycopene into δ-carotene. Accordingly, ε-cyclase activity means the amount of lycopene converted by the ε-cyclase protein, or the amount of δ-carotene formed, in a particular time.

The ε-cyclase activity in genetically modified plants of the invention and in wild-type or reference plants is preferably determined under the following conditions:

the ε-cyclase activity can be determined by the method of Fraser and Sandmann (Biochem. Biophys. Res. Comm. 185(1) (1992) 9-15) *in vitro* if potassium phosphate as buffer (pH 7.6), lycopene as

substrate, stromal protein of paprika, NADP+, NADPH and ATP are added to a defined amount of plant extract.

The ε-cyclase activity in genetically modified plants of the invention and in wild-type and reference plants is particularly preferably determined by the method of Bouvier, d'Harlingue and Camara (Arch Biochem Biophys 346(1) (1997) 53-64): the in vitro assay is carried out in a volume of 0.25 ml. The mixture contains 50 ∝M potassium phosphate (pH 7.6), various amounts of plant extract, 20 nM lycopene, 0.25 mg of paprika chromoplastid stromal protein, 0.2 ∝M NADP+, 0.2 ∝M NADPH and 1 ∝M ATP. NADP/NADPH and ATP are dissolved in 0.01 ml of ethanol with 1 mg of Tween 80 immediately before addition to the incubation medium. After a reaction time of 60 minutes at 30_C, the reaction is stopped by adding chloroform/methanol (2:1). The reaction products extracted into chloroform are analyzed by HPLC.

An alternative assay with radioactive substrate is described in Fraser and Sandmann (Biochem Biophys Res Comm 185(1) (1992) 9-15). A further analytical method is described in Beyer, Kröncke and Nievelstein (J Biol Chem 266(26) (1991) 17072-17078).

In a preferred embodiment of the invention, functional equivalents of the ε-cyclase promoter described by SEQ ID NO: 1, 7 or 8 include all promoters which are located in a plant organism in the 5' direction in front of a genomic sequence which codes for an ε-cyclase having a homology of at least 60%, preferably at least 80%, particularly preferably at least 90%, most preferably at least 95%, with a protein as shown in SEQ ID NO: 10, 12, 14 or 16, where said promoters represent the natural promoter of said genomic sequence.

Functional equivalents of ε-cyclase promoter described by SEQ ID NO: 1, 7 or 8 particularly preferably include all promoters which are located in a plant organism in the 5' direction in front of a genomic sequence which codes for a nucleic acid sequence whose derived cDNA has a homology of at least 60%, preferably at least 80%, particularly preferably at least 90%, most preferably at least 95%, with the nucleic acid sequence as shown in SEQ ID NO: 9, 11, 13 or 15, where said promoters represent the natural promoter of said genomic sequence, and the cDNA codes for an ε-cyclase.

Preferred promoters include a sequence region of least 250 base pairs, preferably at least 500 base pairs, particularly

preferably 1000 base pairs, most preferably at least 2000 base pairs, in the 5' direction calculated from the ATG start codon of said genomic sequences.

Functional equivalents of the ε-cyclase promoter described by SEQ ID NO: 1, 7 or 8 are particularly preferably all promoters which are located in a plant organism in the 5' direction in front of a genomic sequence which codes for an ε-cyclase which comprises at least one of the following sequence motifs:

1. G(G/C)GPAGL(A/S)(V/L)A (SEQ ID NO: 17)
2. (L/I)(N/G/S)RXYG(K/R)(V/L) (SEQ ID NO: 18)
3. MVFMD(Y/W)RD (SEQ ID NO: 19)
4. PTFLY(A/V)M(P/A) (SEQ ID NO: 20)
5. AXMVHP(S/A)TGY(M/S)V(A/V)R (SEQ ID NO: 21)
6. LWPXER(R/K)RQRXFF (SEQ ID NO: 22)

Very particularly preferred functional equivalents of the promoter described by SEQ ID NO: 1, 7 or 8 are those promoters which are located in a plant organism in the 5' direction in front of a genomic sequence which codes for a protein, where said protein includes at least one of the following sequences:

1. the homologous sequence (H1) from Lactuca sativa as shown in SEQ ID NO: 24,
2. the homologous sequences (H2 and H3) from Adonis palaestina as shown in SEQ ID NO: 26 or 28,
3. the homologous sequence (H4) from Arabidopsis thaliana as shown in SEQ ID NO: 30
4. the homologous sequences (H5 and H6) from Citrus x paradisi as shown in SEQ ID NO: 32 or 34
5. the homologous sequence (H7) from Citrus sinensis as shown in SEQ ID NO: 36
6. the homologous sequence (H8) from Spinacea oleracea as shown in SEQ ID NO: 38
7. the homologous sequence (H9) from Solanum tuberosum as shown in SEQ ID NO: 40
8. the homologous sequences (H10 and H11) from Daucus carota as shown in SEQ ID NO: 42 or 44
9. the homologous sequence (H12) from tomato as shown in SEQ ID NO: 46

Most preferred functional equivalents of the promoter described by SEQ ID NO: 1, 7 or 8 are those promoters which are located in a plant organism in the 5' direction in front of a genomic

sequence which codes for a nucleic acid sequence whose derived cDNA includes at least one of the following sequences:

1. the homologous sequence (H1) from Lactuca sativa as shown in SEQ ID NO: 23,
2. the homologous sequences (H2 and H3) from Adonis palaestina as shown in SEQ ID NO: 25 or 27,
3. the homologous sequence (H4) from Arabidopsis thaliana as shown in SEQ ID NO: 29
4. the homologous sequences (H5 and H6) from Citrus x paradisi as shown in SEQ ID NO: 31 or 33
7. the homologous sequence (H7) from Citrus sinensis as shown in SEQ ID NO: 35
5. the homologous sequence (H8) from Spinacea oleracea as shown in SEQ ID NO: 37
6. the homologous sequence (H9) from Solanum tuberosum as shown in SEQ ID NO: 39
8. the homologous sequences (H10 and H11) from Daucus carota as shown in SEQ ID NO: 41 or 43
9. the homologous sequence (H12) from tomato as shown in SEQ ID NO: 45.

Further examples of functionally equivalent promoter sequences employed in the transgenic expression cassettes or transgenic expression vectors of the invention can easily be found for example in various organisms whose genomic sequence is at least partly known, such as, for example, Arabidopsis thaliana, Brassica napus, Nicotiana tabacum, Solanum tuberosum, Helianthus annuus, Linum sativum, followed by homology comparisons in databases.

A further aspect of the invention relates to the use of at least one nucleic acid sequence or of a part thereof in methods for identifying and/or isolating promoters of genes which code for said nucleic acid sequence, where said nucleic acid sequence codes for an amino acid sequence which includes at least one sequence motif as shown in SEQ ID NO: 17, 18, 19, 20, 21 or 22 or a variation indicated for these sequence motifs. Said nucleic acid sequence preferably codes for an amino acid sequence including a sequence as shown in SEQ ID NO: 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44 or 46. Said nucleic acid sequence particularly preferably includes a sequence as shown in SEQ ID NO: 23, 25, 27, 29, 29, 31, 33, 35, 37, 39, 41, 43 or 45. ''Part'' means in relation to the nucleic acid sequence preferably a sequence of at least 10 bases, preferably 15 bases, particularly preferably 20 bases, most preferably 30 bases.

Further included according to the invention are methods for identifying and/or isolating promoters of genes which code for a promoter having specificity for the flower of plants, where at least one nucleic acid sequence or a part thereof is employed in the identification and/or isolation, where said nucleic acid sequence codes for an amino acid sequence which includes at least one sequence motif as shown in SEQ ID NO: 17, 18, 19, 20, 21 or 22 or a variation indicated for these sequence motifs. Said nucleic acid sequence preferably codes for an amino acid sequence including a sequence as shown in SEQ ID NO: 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44 or 46. Said nucleic acid sequence particularly preferably includes a sequence as shown in SEQ ID NO: 23, 25, 27, 29, 29, 31, 33, 35, 37, 39, 41, 43 or 45. ''Part'' means in relation to the nucleic acid sequence preferably a sequence of at least 10 bases, preferably 15 bases, particularly preferably 20 bases, most preferably 30 bases. In a preferred embodiment, the method of the invention is based on the polymerase chain reaction, where said nucleic acid sequence or a part thereof is employed as primer.

Various methods for identifying and isolating, starting from a nucleic acid sequence (e.g. a gene transcript such as, for example, a cDNA), the promoter of the corresponding gene are known to the skilled worker. In principle, all methods for amplifying flanking chromosomal sequences are available for example for this purpose. The two most commonly used methods are inverse PCR (''iPCR''; diagrammatically depicted in Fig. 13) and ''thermal asymmetric interlaced PCR'' (''TAIL PCR''). Also suitable in addition is the method of PCR walkings (Devic et al. (1997) Plant Physiol Biochem 35:331-339).

For the iPCR, genomic DNA of the organism from which the functionally equivalent promoter is to be isolated is completely digested with a given restriction enzyme, and then the individual fragments are religated, i.e. linked to themselves to give a circular molecule, in a diluted mixture. The large number of resulting circular DNA molecules also includes those comprising the known sequence (for example the sequence coding for the homologous protein). Starting from this, the circular molecule can be amplified by PCR using a primer pair where both primers are able to anneal to the known sequence segment. One possible embodiment of the iPCR is reproduced in example 2.

The TAIL-PCR is based on the use of firstly a set of successively truncated highly specific primers which anneal to the known genomic sequence (for example the sequence coding for the

homologous protein), and secondly a set of shorter random primers with a lower melting temperature, so that a less sequence-specific annealing to genomic DNA flanking the known genomic sequence takes place. Annealing of the primers to the DNA to be amplified is possible with such a primer combination to make specific amplification of the desired target sequence possible. One possible embodiment of the TAIL-PCR is reproduced for example in example 2.

A further aspect of the invention relates to methods for preparing a transgenic expression cassette having specificity for the flowers of plants, including the following steps:

I. isolation of a promoter sequence, where at least one nucleic acid sequence or a part thereof is employed in the isolation, where said nucleic acid sequence codes for an amino acid sequence which includes at least one sequence motif as shown in SEQ ID NO: 17, 18, 19, 20, 21 or 22 or a variation indicated for these sequence motifs.

II. functional linkage of said promoter sequence to a further nucleic acid sequence, where said nucleic acid sequence is heterologous in relation to the promoter.

Said nucleic acid sequence preferably codes for an amino acid sequence including a sequence as shown in SEQ ID NO: 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44 or 46. Said nucleic acid sequence particularly preferably includes a sequence as shown in SEQ ID NO: 23, 25, 27, 29, 29, 31, 33, 35, 37, 39, 41, 43 or 45. ''Part'' means in relation to the nucleic acid sequence preferably a sequence of at least 10 bases, preferably 15 bases, particularly preferably 20 bases, most preferably 30 bases. In a preferred embodiment, the method of the invention is based on the polymerase chain reaction, where said nucleic acid sequence or a part thereof is employed as primer. Methods known to the skilled worker, such as, for example, ligation etc., can be employed for the functional linkage (see below).

The level of expression of a functionally equivalent promoter can be both downwards and upwards compared with the promoter found in SEQ ID NO: 1, 7 or 8. Preference is given in this connection to the sequences whose level of expression, measured on the basis of the transcribed mRNA or the subsequently translated protein, under conditions which are otherwise unchanged differs quantitatively by not more than 50%, preferably 25%, particularly preferably 10%, from a comparison value obtained with the

promoters described by SEQ ID NO: 1, 7 or 8. Particularly preferred sequences are those whose level of expression, measured on the basis of the transcribed mRNA or the subsequently translated protein, under conditions which are otherwise unchanged exceeds quantitatively by more than 50%, preferably 100%, particularly preferably 500%, very particularly preferably 1000%, a comparison value obtained with the promoter described by SEQ ID NO: 1, 7 or 8. The preferred comparison value is the level of expression of the mRNAs, naturally expressed from the promoter, of an ε-cyclase or of the protein resulting therefrom. Also preferred as comparison value is the level of expression obtained with any defined nucleic acid sequence, preferably nucleic acid sequences which code for easily quantifiable proteins. Very particular preference is given in this connection to reporter proteins (Schenborn E & Groskreutz D (1999) Mol Biotechnol 13(1):29-44) such as the green fluorescent protein (GFP) (Chiu WL et al. (1996) Curr Biol 6:325-330; Leffel SM et al. (1997) Biotechniques. 23(5):912-8), chloramphenicol acetyltransferase, a luciferase (Millar et al. (1992) Plant Mol Biol Rep 10:324-414) or ß-glucuronidase, very particularly preferably ß-glucuronidase (Jefferson et al. (1987) EMBO J 6:3901-3907).

Functional equivalents also include natural or artificial mutations of the promoter sequence described in SEQ ID NO: 1, 7 or 8. Mutations include substitutions, additions, deletions, inversions or insertions of one or more nucleotide residues. Thus, for example, the present invention also includes nucleotide sequences obtained by modification of the ε-cyclase promoter as shown in SEQ ID NO: 1, 7 or 8. The aim of such a modification may be further localization of the sequence contained therein or, for example, also the insertion or deletion of restriction endonuclease cleavage sites, the deletion of excess DNA or the addition of further sequences, for example further regulatory sequences.

Where insertions, deletions or substitutions, such as, for example, transitions and transversions, are appropriate, it is possible to use techniques known per se, such as in vitro mutagenesis, primer repair, restriction or ligation. Transition means a base-pair exchange of a purine/pyrimidine pair into another purine/pyrimidine pair (e.g. A-T for G-C). Transversion means a base-pair exchange of a purine/pyrimidine pair for a pyrimidine/purine pair (e.g. A-T for T-A). Deletion means removal of one or more base pairs. Insertion means introduction of one or more base pairs.

Complementary ends of the fragments for ligation can be made available by manipulations such as, for example, restriction, chewing back or filling in of overhangs for blunt ends. Analogous results are also obtainable by using the polymerase chain reaction (PCR) using specific oligonucleotide primers.

Homology between two nucleic acids means the identity of the nucleic acid sequence over the complete sequence length in each case, which is calculated by comparison with the aid of the GAP program algorithm (Wisconsin Package Version 10.0, University of Wisconsin, Genetics Computer Group (GCG), Madison, USA), setting the following parameters:

Gap Weight: 12	Length Weight: 4
Average Match: 2.912	Average Mismatch:-2.003

For example, a sequence which has a homology of at least 50% based on nuleic acids with the sequence SEQ ID NO: 1 means a sequence which has a homology of at least 50% on comparison with the sequence SEQ ID NO: 1 by the above program algorithm with the above set of parameters.

Homology between two polypeptides means the identity of the amino acid sequence over the respective sequence length, which is calculated by comparison with the aid of the GAP program algorithm (Wisconsin Package Version 10.0, University of Wisconsin, Genetics Computer Group (GCG), Madison, USA), setting the following parameters:

Gap Weight: 8	Length Weight: 2
Average Match: 2.912	Average Mismatch:-2.003

For example, a sequence having a homology of at least 60% based on protein with the sequence SEQ ID NO: 10 means a sequence which has a homology of at least 60% on comparison with the sequence SEQ ID NO: 10 by the above program algorithm with the above set of parameters.

Functional equivalents also means DNA sequences which hybridize under standard conditions with the nucleic acid sequence coding for the ε-cyclase promoter as shown in SEQ ID NO: 1, 7 or 8, or with the nucleic acid sequences complementary thereto, and which have substantially the same promoter properties. The term standard hybridization conditions is to be understood broadly and

means both stringent and less stringent hybridization conditions. Such hybridization conditions are described inter alia in Sambrook J, Fritsch EF, Maniatis T et al., in Molecular Cloning - A Laboratory Manual, 2nd edition, Cold Spring Harbor Laboratory Press, 1989, pages 9.31-9.57 or in Current Protocols in Molecular Biology, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6.

For example, the conditions during the washing step can be selected from the range of conditions limited by those of low stringency (with approximately 2X SSC at 50_C) and of high stringency (with approximately 0.2X SSC at 50_C, preferably at 65_C) (20X SSC: 0.3 M sodium citrate, 3 M NaCl, pH 7.0). In addition, the temperature during the washing step can be raised from low-stringency conditions at room temperature, approximately 22_C, to more stringent conditions at approximately 65_C. Both parameters, the salt concentration and the temperature, can be varied simultaneously, and it is also possible for one of the two parameters to be kept constant and only the other to be varied. It is also possible to employ denaturing agents such as, for example, formamide or SDS during the hybridization. Hybridization in the presence of 50% formamide is preferably carried out at 42_C. Some exemplary conditions for hybridization and washing steps are given below:

(1) Hybridization conditions with for example

a) 4X SSC at 65_C, or
b) 6X SSC, 0.5% SDS, 100 αg/ml denatured fragmented salmon sperm DNA at 65_C, or
c) 4X SSC, 50% formamide, at 42_C, or
d) 2X or 4X SSC at 50_C (low-stringency condition), or
e) 2X or 4X SSC, 30 to 40% formamide at 42_C (low-stringency condition), or
f) 6x SSC at 45_C, or,
g) 0.05 M sodium phosphate buffer pH 7.0, 2 αM EDTA, 1% BSA and 7% SDS.

(2) Washing steps with for example

a) 0.1X SSC at 65_C, or
b) 0.1X SSC, 0.5% SDS at 68_C, or
c) 0.1X SSC, 0.5% SDS, 50% formamide at 42_C, or
d) 0.2X SSC, 0.1% SDS at 42_C, or
e) 2X SSC at 65_C (low-stringency condition), or
f) 40 αM sodium phosphate buffer pH 7.0, 1% SDS, 2 αM EDTA.

Methods for preparing functional equivalents of the invention preferably include the introduction of mutations into the ε-cyclase promoter as shown in SEQ ID NO: 1, 7 or 8. Mutagenesis may be random, in which case the mutagenized sequences are subsequently screened for their properties by a trial and error procedure. Particularly advantageous selection criteria include for example the level of the resulting expression of the introduced nucleic acid sequence in a flower tissue.

Methods for mutagenesis of nucleic acid sequences are known to the skilled worker and include by way of example the use of oligonucleotides with one or more mutations compared with the region to be mutated (e.g. in a site-specific mutagenesis). Primers with approximately 15 to approximately 75 nucleotides or more are typically employed, with preferably about 10 to about 25 or more nucleotide residues being located on both sides of the sequence to be modified. Details and procedure for said mutagenesis methods are familiar to the skilled worker (Kunkel et al. (1987) Methods Enzymol 154:367-382; Tomic et al. (1990) Nucl Acids Res 12:1656; Upender et al. (1995) Biotechniques 18(1):29-30; US 4,237,224). A mutagenesis can also be achieved by treating for example transgenic expression vectors comprising one of the nucleic acid sequences of the invention with mutagenizing agents such as hydroxylamine.

An alternative possibility is to delete nonessential sequences of a promoter of the invention without significantly impairing the essential properties mentioned. Such deletion variants represent functionally equivalent fragments to the promoters described by SEQ ID NO: 1, 7 or 8 or to functional equivalents thereof. Localization of the promoter sequence to particular essential regulatory regions can be carried out for example with the aid of search routine to search for promoter elements. Particular promoter elements are often present in increased numbers in the regions relevant for promoter activity. This analysis can be carried out for example with computer programs such as the PLACE program (''Plant Cis-acting Regulatory DNA Elements''; Higo K et al. (1999) Nucl Acids Res 27(1): 297-300), the BIOBASE database ''Transfac'' (Biologische Datenbanken GmbH, Braunschweig; Wingender E et al. (2001) Nucleic Acids Res 29(1):281-3) or the PlantCARE database (Lescot M et al. (2002) Nucleic Acids Res 30(1):325-7).

The functionally equivalent fragments of one of the promoters of the invention - for example of the ε-cyclase promoters described by SEQ ID NO: 1, 7 or 8 - preferably include at least 200 base

pairs, very particularly preferably at least 500 base pairs, most preferably at least 1000 base pairs of the 3' end of the respective promoter of the invention - for example the promoters described by SEQ ID NO: 1, 7 or 8 - the length being calculated from the translation start (''ATG'' codon) upstream in the 5' direction.

Further functionally equivalent fragments may be generated for example by deleting any 5'-untranslated regions still present. For this purpose, the start of transcription of the corresponding genes can be determined by methods familiar to the skilled worker (such as, for example, 5'-RACE), and the 5'-untranslated regions can be deleted by PCR-mediated methods or endonuclease digestion. Thus, for example, the 5'-untranslated regions included in the promoters shown in SEQ ID NO: 7 or 8 can be deleted without the promoter losing its essential functionality. Corresponding deletion variants are expressly included as functional equivalents.

In transgenic expression cassettes of the invention, at least one of the promoters of the invention (e.g. described by SEQ ID NO: 1, 7 or 8) is functionally linked to at least one nucleic acid sequence to be expressed transgenically.

A functional linkage means, for example, the sequential arrangement of one of the promoters of the invention (e.g. described by SEQ ID NO: 1, 7 or 8) with a nucleic acid sequence to be expressed transgenically and, where appropriate, further genetic control sequences such as, for example, a terminator or a polyadenylation sequence in such a way that the promoter is able to fulfill its function in the transgenic expression of the nucleic acid sequence under suitable conditions, and expression of the nucleic acid sequence (i.e. transcription and, where appropriate, translation) takes place. ''Suitable conditions'' means in this connection preferably the presence of the expression cassette in a plant cell, preferably a plant cell included in a flower of the plant.

Arrangements in which the nucleic acid sequence to be expressed transgenically is positioned behind one of the promoters of the invention (e.g. described by SEQ ID NO: 1, 7 or 8), so that the two sequences are covalently connected together, are preferred. In this connection, the distance between the promoter sequence and the nucleic acid sequence to be expressed transgenically is preferably fewer than 200 base pairs, particularly preferably

less than 100 base pairs, very particularly preferably less than 50 base pairs.

Production of a functional linkage and production of a transgenic expression construct can be achieved by means of conventional recombination and cloning techniques as described for example in Maniatis T, Fritsch EF and Sambrook J (1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor (NY), in Silhavy TJ, Berman ML and Enquist LW (1984) Experiments with Gene Fusions, Cold Spring Harbor Laboratory, Cold Spring Harbor (NY) and in Ausubel FM et al. (1987) Current Protocols in Molecular Biology, Greene Publishing Assoc. and Wiley Interscience. However, further sequences which have for example the function of a linker with particular restriction enzyme cleavage sites or of a signal peptide may also be positioned between the two sequences. Insertion of sequences may also lead to expression of fusion proteins. It is possible and preferred for the transgenic expression construct, consisting of a linkage of promoter and nucleic acid sequence to be expressed, to be integrated into a vector and be inserted into a plant genome for example by transformation.

However, an expression cassette also means constructions in which one of the promoters of the invention (e.g. described by SEQ ID NO: 1, 7 or 8) is, without necessarily having been functionally linked beforehand to a nucleic acid sequence to be expressed, introduced into a host genome, for example by targeted homologous recombination or random insertion, there undertakes regulatory control over endogenous nucleic acid sequences then functionally linked thereto, and controls the transgenic expression thereof. Insertion of the promoter - for example by a homologous recombination - in front of a nucleic acid coding for a particular polypeptide results in an expression cassette of the invention which controls the targeted expression of the particular polypeptide in the flower of plants. It is also possible for example for the natural promoter of an endogenous gene to be replaced by one of the promoters of the invention (e.g. described by SEQ ID NO: 1, 7 or 8), and for the expression behavior of the endogenous gene to be modified.

A further possibility is also for the promoter to be inserted in such a way that antisense RNA or a double-stranded RNA (e.g. in the form of an inverted repeat) is expressed to give the nucleic acid coding for a particular polypeptide. In this way, expression of the particular polypeptide in the flower of plants is selectively downregulated or switched off.

It is also possible analogously for a nucleic acid sequence which is to be expressed transgenically to be placed - for example by homologous recombination - behind the sequence which codes for one of the promoters of the invention (e.g. described by SEQ ID NO: 1, 7 or 8), and which is located in its natural chromosomal context, so as to result in an expression cassette of the invention which controls the expression of the nucleic acid sequence to be expressed transgenically in the flower of plants.

The transgenic expression cassettes of the invention may include further genetic control sequences. The term genetic control sequences is to be understood broadly and means all sequences having an influence on the coming into existence or the function of a transgenic expression cassette of the invention. Genetic control sequences modify for example the transcription and translation in prokaryotic or eukaryotic organisms. The transgenic expression cassettes of the invention preferably include as additional genetic control sequence a terminator sequence 3'-downstream from the particular nucleic acid sequence to be expressed transgenically, and where appropriate further customary regulatory elements, in each case functionally linked to the nucleic acid sequence to be expressed transgenically.

Genetic control sequences also include further promoters, promoter elements or minimal promoters able to modify the expression-controlling properties. It is thus possible for example through genetic control sequences for tissue-specific expression to take place additionally in dependence on particular stress factors. Corresponding elements are described for example for water stress, abscisic acid (Lam E and Chua NH, J Biol Chem 1991; 266(26):17131-17135) and heat stress (Schoffl F et al. (1989) Mol Gen Genetics 217(2-3):246-53).

A further possibility is for further promoters which make transgenic expression possible in further plant tissues or in other organisms such as, for example, *E.coli* bacteria to be functionally linked to the nucleic acid sequence to be expressed. Suitable promoters are in principle all promoters functional in plants. Promoters functional in plants means in principle every promoter able to control the expression of genes, in particular foreign genes, in plants or plant parts, cells, tissues, cultures. It is moreover possible for expression to be for example constitutive, inducible or development-dependent. Preference is given to constitutive promoters, tissue-specific promoters, development-dependent promoters, chemically inducible,

stress-inducible or pathogen-inducible promoters. Corresponding promoters are generally known to the skilled worker.

Further advantageous control sequences are to be found for example in the promoters of gram-positive bacteria such as amy and SPO2 or in the yeast or fungal promoters ADC1, MFα, AC, P-60, CYC1, GAPDH, TEF, rp28, ADH.

It is possible in principle for all natural promoters with their regulatory sequences like those mentioned above to be used for the method of the invention. It is additionally also possible for synthetic promoters to be used advantageously.

Genetic control sequences further include also the 5'-untranslated regions, introns or noncoding 3' region of genes such as, for example, the actin-1 intron, or the Adh1-S introns 1, 2 and 6 (generally: The Maize Handbook, Chapter 116, Freeling and Walbot, Eds., Springer, New York (1994)), preferably the genes with the gene locus At2g46720, At3g01980 and At1g63140 from Arabidopsis thaliana. It is possible to show that such regions may have a significant function in regulating gene expression. Thus, it has been shown that 5'-untranslated sequences are able to enhance the transient expression of heterologous genes. Examples of translation enhancers which may be mentioned are the 5' leader sequence from tobacco mosaic virus (Gallie et al. (1987) Nucl Acids Res 15:8693-8711) and the like. They may in addition promote tissue specificity (Rouster J et al. (1998) Plant J 15:435-440). The nucleic acid sequences indicated in SEQ ID NO: 2, 7 or 8 in each case represent the promoter region and the 5'-untranslated regions up to the ATG start codon of the respective genes.

The transgenic expression construct may advantageously comprise one or more so-called enhancer sequences functionally linked to the promoter, which make increased transgenic expression of the nucleic acid sequence possible. Additional advantageous sequences can also be inserted at the 3' end of the nucleic acid sequences to be expressed transgenically, such as further regulatory elements or terminators. The nucleic acid sequences to be expressed transgenically may be present in one or more copies in the gene construct.

Polyadenylation signals suitable as control sequences are plant polyadenylation signals, preferably those which are essentially T-DNA polyadenylation signals from *Agrobacterium tumefaciens*. Examples of particularly suitable terminator sequences are the

OCS (octopine synthase) terminator and the NOS (nopaline synthase) terminator.

Control sequences additionally mean those which make homologous recombination or insertion into the genome of a host organism possible or allow deletion from the genome. In homologous recombination for example the coding sequence of a particular endogenous gene can be specifically replaced by a sequence coding for a dsRNA. Methods such as cre/lox technology permit tissue-specific, and in some circumstances inducible, deletion of the transgenic expression constant from the genome of the host organism (Sauer B (1998) Methods 14(4):381-92). In this case, particular flanking sequences are attached to the target gene (lox sequences) and make later deletion by means of cre recombinase possible.

A transgenic expression cassette and/or the transgenic expression vectors derived therefrom may comprise further functional elements. The term functional element is to be understood broadly and means all elements which have an influence on the production, replication or function of the transgenic expression constructs of the invention, of the transgenic expression vectors or of the transgenic organisms. Non-restrictive examples which may be mentioned are:

a) Selection markers which confer resistance to biocides such as metabolism inhibitors (e.g. 2-deoxyglucose 6-phosphate; WO 98/45456), antibiotics (e.g. kanamycin, G 418, bleomycin, hygromycin) or - preferably - herbicides (e.g. phosphinothricin). Examples of selection markers which may be mentioned are: phosphinothricin acetyltransferases (bar and pat gene), which inactivate glutamine synthase inhibitors, 5-enolpyruvylshikimate-3-phosphate synthases (EPSP synthase genes) which confer resistance to glyphosate (N-(phosphonomethyl)glycine), glyphosate-degrading enzymes (gox gene product; glyphosate oxidoreductase), dehalogenases which for example inactivate dalapon (deh gene product), sulfonylurea- and imidazolinone-inactivating acetolactate synthases, and nitrilases which for example degrade bromoxynil (bxn gene product), the aasa gene product which confers resistance to the antibiotic spectinomycin, streptomycin phosphotransferases (SPT) which ensure resistance to streptomycin, neomycin phosphotransferases (NPTII) which confer resistance to kanamycin or geneticin, the hygromycin phosphotransferases (HPT) which mediate resistance to hygromycin, the acetolactate synthases (ALS)

which confer resistance to sulfonylurea herbicides (e.g. mutated ALS variants with, for example, the S4 and/or Hra mutation).

b) Reporter genes which code for easily quantifiable proteins and ensure via an intrinsic color or enzymic activity an assessment of the transformation efficiency or of the location or timing of expression. Very particular preference is given in this connection to reporter proteins (Schenborn E, Groskreutz D. Mol Biotechnol. 1999; 13(1):29-44) such as the green fluorescent protein (GFP) (Sheen et al.(1995) Plant Journal 8(5):777-784), the chloramphenicol acetyltransferase, a luciferase (Ow et al. (1986) Science 234:856-859), the aequorin gene (Prasher et al. (1985) Biochem Biophys Res Commun 126(3):1259-1268), the β-galactosidase, with very particular preference for ß-glucuronidase (Jefferson et al. (1987) EMBO J 6:3901-3907).

c) Origins of replication which ensure replication of the transgenic expression constructs or transgenic expression vectors of the invention in, for example, E. coli. Examples which may be mentioned are ORI (origin of DNA replication), the pBR322 ori or the P15A ori (Sambrook et al.: Molecular Cloning. A Laboratory Manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989).

d) Elements which are necessary for agrobacterium-mediated plant transformation, such as, for example, the right or left border of the T-DNA or the vir region.

''Introduction'' includes for the purposes of the invention all methods suitable for introducing a nucleic acid sequence (for example an expression cassette of the invention) directly or indirectly into an organism (e.g. a plant) or a cell, compartment, tissue, organ or propagation material (e.g. seeds or fruits) thereof, or for generating such therein. Direct and indirect methods are included. The introduction can lead to a temporary (transient) presence of said nucleic acid sequence or else to a permanent (stable) presence. Introduction includes for example methods such as transfection, transduction or transformation. The organisms used in the methods are grown or cultured, depending on the host organism, in the manner known to the skilled worker.

Introduction of a transgenic expression cassette of the invention into an organism or cells, tissues, organs, parts or seeds

thereof (preferably into plants or plant cells, tissues, organs, parts or seeds) can advantageously be achieved by use of vectors comprising the transgenic expression cassettes. Vectors may be for example plasmids, cosmids, phages, viruses or else agrobacteria. The transgenic expression cassettes can be inserted into the vector (preferably a plasmid vector) via a suitable restriction cleavage site. The resulting vector can be firstly introduced and amplified in E. coli. Correctly transformed E. coli are selected and cultured, and the recombinant vector is isolated by methods familiar to the skilled worker. Restriction analysis and sequencing can be used to check the cloning step. Preferred vectors are those making stable integration of the expression cassette into the host genome possible.

Production of a transformed organism (or of a transformed cell or tissue) requires introduction of the appropriate DNA (e.g. the expression vector) or RNA into the appropriate host cell. A large number of methods is available for this process, which is referred to as transformation (or transduction or transfection) (Keown et al. (1990) Methods in Enzymology 185:527-537). Thus, the DNA or RNA can for example be introduced directly by microinjection or by bombardment with DNA-coated microparticles. The cell can also be permeabilized chemically, for example with polyethylene glycol, so that the DNA is able to enter the cell by diffusion. The DNA introduction can also take place by protoplast fusion with other DNA-containing units such as minicells, cells, lysosomes or liposomes. Electroporation is another suitable method for introducing DNA, in which the cells are reversibly permeabilized by an electrical impulse. Corresponding methods are described (for example in Bilang et al. (1991) Gene 100:247-250; Scheid et al. (1991) Mol Gen Genet 228:104-112; Guerche et al. (1987) Plant Science 52:111-116; Neuhause et al. (1987) Theor Appl Genet 75:30-36; Klein et al. (1987) Nature 327:70-73; Howell et al. (1980) Science 208:1265; Horsch et al.(1985) Science 227:1229-1231; DeBlock et al. (1989) Plant Physiology 91:694-701; Methods for Plant Molecular Biology (Weissbach and Weissbach, eds.) Academic Press Inc. (1988); and Methods in Plant Molecular Biology (Schuler and Zielinski, eds.) Academic Press Inc. (1989)).

Vectors preferred for expression in E. coli are pQE70, pQE60 and pQE-9 (QIAGEN, Inc.); pBluescript vectors, Phagescript vectors, pNH8A, pNH16a, pNH18A, pNH46A (Stratagene Cloning Systems, Inc.); ptrc99a, pKK223-3, pKK233-3, pDR540, pRIT5 (Pharmacia Biotech, Inc.).

Preferred vectors for expression in mammalian cells include pWLNE0, pSV2CAT, pOG44, pXT1 and pSG (Stratagene Inc.); pSVK3, pBPV, pMSG and pSVL (Pharmacia Biotech, Inc.). Inducible vectors which may be mentioned are pTet-tTak, pTet-Splice, pcDNA4/TO, pcDNA4/TO /LacZ, pcDNA6/TR, pcDNA4/TO/Myc-His/LacZ, pcDNA4/TO/Myc-His A, pcDNA4/TO/Myc-His B, pcDNA4/TO/Myc-His C, pVgRXR (Invitrogen, Inc.) or the pMAM series (Clontech, Inc.; GenBank Accession No: U02443). These themselves provide the inducible regulatory control element for example for a chemically inducible expression.

Vectors for expression in yeast include for example pYES2, pYD1, pTEF1/Zeo, pYES2/GS, pPICZ, pGAPZ, pGAPZalph, pPIC9, pPIC3.5, PHIL-D2, PHIL-S1, pPIC3SK, pPIC9K, and PA0815 (Invitrogen, Inc.).

Cloning vectors and techniques for genetic manipulation of ciliates and algae are known to the skilled worker (WO 98/01572; Falciatore et al. (1999) Marine Biotechnology 1(3):239-251; Dunahay et al. (1995) J Phycol 31:10004-1012).

The methods to be used in principle for the transformation of animal cells or of yeast cells are similar to those for "direct" transformation of plant cells. Methods such as calcium phosphate or liposome-mediated transformation or else electroporation are preferred in particular.

Various methods and vectors for inserting genes into the genome of plants and for regenerating plants from plant tissues or plant cells are known (Plant Molecular Biology and Biotechnology (CRC Press, Boca Raton, Florida), Chapter 6/7, pp. 71-119 (1993); White FF (1993) Vectors for Gene Transfer in Higher Plants; in: Transgenic Plants, Vol. 1, Engineering and Utilization, editors: Kung and Wu R, Academic Press, 15-38; Jenes B et al. (1993) Techniques for Gene Transfer, in: Transgenic Plants, Vol. 1, Engineering and Utilization, editors: Kung and R. Wu, Academic Press, pp. 128-143; Potrykus (1991) Annu Rev Plant Physiol Plant Molec Biol 42:205-225; Halford NG, Shewry PR (2000) Br Med Bull 56(1):62-73). Those mentioned above are included, for example. In the case of plants, the described methods for the transformation and regeneration of plants from plant tissues or plant cells are used for transient or stable transformation. Suitable methods are, in particular, protoplast transformation by polyethylene glycol-induced DNA uptake, calcium phosphate-mediated transformation, DEAE-dextran-mediated transformation, liposome-mediated transformation (Freeman et al. (1984) Plant Cell Physiol. 29:1353ff; US 4,536,475), biolistic methods with the

gene gun (''particle bombardment'' method; US 5,100,792; EP-A 0 444 882; EP-A 0 434 616; Fromm ME et al. (1990) Bio/Technology 8(9):833-9; Gordon-Kamm et al. (1990) Plant Cell 2:603), electroporation, incubation of dry embryos in DNA-containing solution, electroporation (EP-A 290 395, WO 87/06614), microinjection (WO 92/09696, WO 94/00583, EP-A 0 331 083, EP-A 0 175 966) or other methods of direct DNA introduction (DE 4 005 152, WO 90/12096, US 4,684,611). Physical methods of DNA introduction into plant cells are surveyed in Oard (1991) Biotech Adv 9:1-11.

In the case of these ''direct'' transformation methods, no particular requirements need be met by the plasmid used. Simple plasmids such as those of the pUC series, pBR322, M13mp series, pACYC184 etc. can be used. If complete plants are to be regenerated from the transformed cells, it is necessary for an additional selectable marker gene to be present on the plasmid.

Besides these ''direct'' transformation techniques, it is also possible to carry out a transformation by bacterial infection using agrobacterium (e.g. EP 0 116 718), viral infection using viral vectors (EP 0 067 553; US 4,407,956; WO 95/34668; WO 93/03161) or using pollen (EP 0 270 356; WO 85/01856; US 4,684,611).

The transformation is preferably effected using agrobacteria which comprise disarmed Ti plasmid vectors, utilizing their natural ability to transfer genes to plants (EP-A 0 270 355; EP-A 0 116 718).

Agrobacterium transformation is widely used for the transformation of dicotyledons, but is also increasingly being applied to monocotyledons (Toriyama et al. (1988) Bio/Technology 6: 1072-1074; Zhang et al. (1988) Plant Cell Rep 7:379-384; Zhang et al. (1988) Theor Appl Genet 76:835-840; Shimamoto et al. (1989) Nature 338:274-276; Datta et al. (1990) Bio/Technology 8: 736-740; Christou et al. (1991) Bio/Technology 9:957-962; Peng et al. (1991) International Rice Research Institute, Manila, Philippines 563-574; Cao et al. (1992) Plant Cell Rep 11:585-591; Li et al. (1993) Plant Cell Rep 12:250-255; Rathore et al. (1993) Plant Mol Biol 21:871-884; Fromm et al. (1990) Bio/Technology 8:833-839; Gordon-Kamm et al. (1990) Plant Cell 2:603-618; D'Halluin et al. (1992) Plant Cell 4:1495-1505; Walters et al. (1992) Plant Mol Biol 18:189-200; Koziel et al. (1993) Biotechnology 11:194-200; Vasil IK (1994) Plant Mol Biol 25:925-937; Weeks et al. (1993) Plant Physiol 102:1077-1084; Somers et

al. (1992) Bio/Technology 10:1589-1594; WO 92/14828; Hiei et al. (1994) Plant J 6:271-282).

The strains mostly used for agrobacterium transformation, Agrobacterium tumefaciens or Agrobacterium rhizogenes comprise a plasmid (Ti or Ri plasmid) which is transferred to the plant after agrobacterium infection. Part of this plasmid, called T-DNA (transferred DNA), is integrated into the genome of the plant cell. Alternatively, binary vectors (mini-Ti plasmids) can also be transferred into plants and integrated in the genome thereof by agrobacterium.

The use of Agrobacterium tumefaciens for the transformation of plants using tissue culture explants is described (inter alia Horsch RB et al. (1985) Science 225:1229ff.; Fraley et al. (1983) Proc Natl Acad Sci USA 80: 4803-4807; Bevans et al. (1983) Nature 304:184-187). Many Agrobacterium tumefaciens strains are able to transfer genetic material - for example the expression cassettes of the invention - such as, for example, the strains EHA101[pEHA101], EHA105[pEHA105], LBA4404[pAL4404], C58C1[pMP90] and C58C1[pGV2260] (Hood et al. (1993) Transgenic Res 2:208-218; Hoekema et al. (1983) Nature 303:179-181; Koncz and Schell (1986) Gen Genet 204:383-396; Deblaere et al. (1985) Nucl Acids Res 13: 4777-4788).

On use of agrobacteria, the expression cassette must be integrated into specific plasmids either into a shuttle or intermediate vector or into a binary vector. Binary vectors able to replicate both in E. coli and in agrobacterium are preferably used. They normally comprise a selection marker gene and a linker or polylinker, flanked by the right and left T-DNA border sequence. They can be transformed directly into agrobacterium (Holsters et al. (1978) Mol Gen Genet 163:181-187). The agrobacterium acting as host organsim in this case should already comprise a plasmid having the vir region. This is necessary for transfer of the T-DNA into the plant cell. An agrobacterium transformed in this way can be used to transform plant cells. The use of T-DNA for transforming plant cells has been intensively investigated and described (EP-A 0 120 516; Hoekema, In: The Binary Plant Vector System, Offsetdrukkerij Kanters B.V., Alblasserdam, Chapter V; An et al. (1985) EMBO J 4:277-287). Various binary vectors are known, and some of them are commercially available, such as, for example, pBI101.2 or pBIN19 (Clontech Laboratories, Inc. USA; Bevan et al.(1984) Nucl Acids Res 12:8711), pBinAR, pPZP200 or pPTV.

Agrobacteria transformed with such a vector can then be used in a known manner for transforming plants, especially crop plants such as, for example, oilseed rape, by for example bathing wounded leaves or pieces of leaf in a solution of agrobacteria and then cultivating in suitable media. Transformation of plants by agrobacteria is described (White FF (1993) Vectors for Gene Transfer in Higher Plants; in Transgenic Plants, Vol. 1, Engineering and Utilization, edited by SD Kung and R Wu, Academic Press, pp. 15-38; Jenes B et al.(1993) Techniques for Gene Transfer, in: Transgenic Plants, Vol. 1, Engineering and Utilization, edited by S.D. Kung and R. Wu, Academic Press, pp. 128-143; Potrykus (1991) Annu Rev Plant Physiol Plant Molec Biol 42:205-225). Transgenic plants which have integrated the expression systems of the invention described above can be regenerated in a known manner from the transformed cells of the wounded leaves or pieces of leaf.

Stably transformed cells (i.e. those which have integrated the introduced DNA into the DNA of the host cell) can be selected from untransformed ones if a selectable marker is a constituent of the introduced DNA. Any gene able to confer a resistance to a biocide (e.g. an antibiotic or herbicide, see above) can act as marker, for example. Transformed cells which express such a marker gene are able to survive in the presence of concentrations of a corresponding biocide which kill an untransformed wild type. The selection marker permits the selection of transformed cells from untransformed ones (McCormick et al. (1986) Plant Cell Reports 5:81-84). The resulting plants can be grown and crossed in the usual way. Two or more generations should be cultivated in order to ensure that the genomic integration is stable and heritable.

As soon as a transformed plant cell has been produced, it is possible to obtain a complete plant by using methods known to the skilled worker. These entail, for example, starting from callus cultures, single cells (e.g. protoplasts) or leaf disks (Vasil et al. (1984) Cell Culture and Somatic Cell Genetics of Plants, Vol I, II and III, Laboratory Procedures and Their Applications, Academic Press; Weissbach and Weissbach (1989) Methods for Plant Molecular Biology, Academic Press). The formation of shoot and root from these still undifferentiated callus cell masses can be induced in a known manner. The resulting shoots can be planted out and grown. Corresponding methods are described (Fennell et al. (1992) Plant Cell Rep. 11: 567-570; Stoeger et al (1995) Plant Cell Rep. 14:273-278; Jahne et al. (1994) Theor Appl Genet 89:525-533).

The effectiveness of expression of the transgenically expressed nucleic acids can be estimated for example *in vitro* by shoot-meristem propagation using one of the selection methods described above. In addition, a change in the type and level of expression of a target gene, and the effect on the phenotype of the plant can be tested on test plants in glasshouse tests.

A further aspect of the invention relates to transgenic organisms transformed with at least one expression cassette of the invention or one vector of the invention, and cells, cell cultures, tissues, parts - such as, for example, in the case of plant organisms leaves, roots etc. - or propagation material derived from such organisms.

By organism, starting or host organisms are meant prokaryotic or eukaryotic organisms such as, for example, microorganisms or plant organisms. Preferred microorganisms are bacteria, yeasts, algae or fungi.

Preferred bacteria are bacteria of the genus Escherichia, Erwinia, Agrobacterium, Flavobacterium, Alcaligenes, Pseudomonas, Bacillus or cyanobacteria, for example of the genus Synechocystis and further bacterial genera described in Brock Biology of Microorganisms Eighth Edition on pages A-8, A-9, A10 and A11.

Microorganisms which are particularly preferred are those able to infect plants and thus transfer the constructs of the invention. Preferred microorganisms are those of the genus Agrobacterium and especially of the species Agrobacterium tumefaciens. Particularly preferred microorganisms are those able to produce toxins (e.g. botulinum toxin), pigments (e.g. carotenoids or flavonoids), antibiotics (e.g. penicillin), phenylpropanoids (e.g. tocopherol), polyunsaturated fatty acids (e.g. arachidonic acid) or vitamins (e.g. vitamin B12).

Preferred yeasts are Candida, Saccharomyces, Hansenula or Pichia.

Preferred fungi are Aspergillus, Trichoderma, Ashbya, Neurospora, Fusarium, Beauveria or further fungi described in Indian Chem Engr. Section B. Vol 37, No. 1,2 (1995) on page 15, table 6.

Host or starting organisms preferred as transgenic organisms are in particular plant organisms.

''Plant organism of cells derived therefrom'' means in general every cell, tissue, part or propagation material (such as seeds or fruits) of an organism capable of photosynthesis. Included for the purposes of the invention are all genera and species of higher and lower plants of the plant kingdom. Annual, perennial, monocotyledonous and dicotyledonous plants are preferred.

''Plant'' means for the purposes of the invention all genera and species of higher and lower plants of the plant kingdom. The term includes the mature plants, seeds, shoots and seedlings, and parts derived therefrom, propagation material (for example tubers, seeds or fruits), plant organs, tissues, protoplasts, callus and other cultures, for example cell or callus cultures, and all other types of groupings of plant cells to functional or structural units. Mature plants means plants at any stage of development beyond seedling. Seedling means a young, immature plant at an early stage of development.

Plant organisms for the purposes of the invention are additionally further photosynthetically active organisms such as, for example, algae, cyanobacteria and mosses. Preferred algae are green algae, such as, for example, algae of the genus Haematococcus, Phaedactylum tricornatum, Volvox or Dunaliella. Synechocystis, Chlamydomonas and Scenedesmus are particularly preferred.

Particularly preferred for the purposes of the method of the invention are plant organisms selected from the group of flowering plants (Phylum Anthophyta ''angiosperms''). All annual and perennial, monocotyledonous and dicotyledonous plants are included. The plant is preferably selected from the following plant families: Amaranthaceae, Asteraceae, Brassicaceae, Caryophyllaceae, Chenopodiaceae, Compositae, Cruciferae, Cucurbitaceae, Labiatae, Leguminosae, Papilionoideae, Liliaceae, Linaceae, Malvaceae, Rosaceae, Rubiaceae, Saxifragaceae, Scrophulariaceae, Solanaceae, Sterculiaceae, Tetragoniaceae, Theaceae and Umbelliferae.

The invention is very particularly preferably applied to dicotyledonous plant organisms. Preferred dicotyledonous plants are in particular selected from the dicotyledonous crop plants such as, for example the following

1) Category: Dicotyledonae (dicotyledons). Preferred families:

- Aceraceae (maples)

- Cactaceae (cacti)

- Rosaceae (roses, apples, almonds, strawberries)

- Salicaceae (willows)

- Asteraceae (compositae) especially the genus Lactuca, very especially the species sativa (lettuce), and sunflower, dandelion, Tagetes or Calendula and many others,

- Cruciferae (Brassicaceae), especially the genus Brassica, very especially the species napus (oilseed rape), campestris (beet), oleracea (e.g. cabbage, cauliflower or broccoli and other brassica species); and of the genus Arabidopsis, very especially the species thaliana, and cress, radish, canola and many others,

- Cucurbitaceae such as melon, pumpkin, cucumber or zucchini and many others,

- Leguminosae (Fabaceae) especially the genus Glycine, very especially the species max (soybean), soya and alfalfa, pea, beans, lupin or peanut and many others,

- Malvaceae, especially mallow, cotton, edible marshmallow, hibiscus and many others,

- Rubiaceae, preferably of the subclass Lamiidae such as, for example, Coffea arabica or Coffea liberica (coffee bush) and many others,

- Solanaceae, especially the genus Lycopersicon, very especially the species esculentum (tomato) and the genus Solanum, very especially the species tuberosum (potato) and melongena (eggplant) and the genus Capsicum, very especially the species annuum (paprika), and tobacco, petunia and many others,

- Sterculiaceae, preferably of the subclass Dilleniidae such as, for example, Theobroma cacao (cocoa plant) and many others,

- Theaceae, preferably of the subclass Dilleniidae such as, for example, Camellia sinensis or Thea sinensis (tea bush) and many others,

- Umbelliferae (Apiaceae), especially the genus Daucus (very especially the species carota (carrot)), Apium (very especially the species graveolens dulce (celeriac)), and parsley and many others;

and flax, hemp, spinach, carrot, sugarbeet and the various tree, nut and vine species, especially bannana and kiwi fruit.

However, in addition, monocotyledonous plants are also suitable. These are preferably selected from the monocotylendonous crop plants such as, for example the families

- Arecaceae (palms)
- Bromeliaceae (pineapple, spanish moss)
- Cyperaceae (sedges)
- Liliaceae (lilies, tulips, hyacinths, onions, garlic)
- Orchidaceae (orchids)
- Poaceae (grasses, bamboos, corn, sugarcane, wheat)
- Iridaceae (buckwheat, gladioli, crocuses)

Very particular preference is given to Gramineae such as rice, corn, wheat or other cereal species such as barley, millet, rye, triticale or oats, and the sugarcane, and all species of grasses.

Within the framework of the expression cassette of the invention, expression of a particular nucleic acid may, through a promoter having specificity for the flower of plants, lead to the formation of sense RNA, antisense RNA or double-stranded RNA in the form of an inverted repeat (dsRNAi). The sense RNA can subsequently be translated into particular polypeptides. It is possible with the antisense RNA and dsRNAi to down regulate the expression of particular genes.

The method of gene regulation by means of double-stranded RNA (''double-stranded RNA interference''; dsRNAi) has been described in animal and plant organisms many times (e.g. Matzke MA et al. (2000) Plant Mol Biol 43:401-415; Fire A et al (1998) Nature 391:806-811; WO 99/32619; WO 99/53050; WO 00/68374; WO 00/44914; WO 00/44895; WO 00/49035; WO 00/63364). Express reference is made to the processes and methods described in the citations indicated.

The specificity of the expression constructs and vectors of the invention for flowers of plants is particularly advantageous. The flower has the function in attracting beneficial insects through incorporation of pigments or synthesis of volatile chemicals.

The natural defense mechanisms of the plant, for example against pathogens, are often inadequate. Introduction of foreign genes from plants, animals or microbial sources may enhance the defenses. Examples are protection against insect damage to tobacco through expression of the Bacillus thuringiensis endotoxin (Vaeck et al. (1987) Nature 328:33-37) or protection of tobacco from fungal attack through expression of a chitinase from beans (Broglie et al. (1991) Science 254:1194-1197).

Cold spells during the flowering period lead to considerable crop losses every year. Targeted expression of protective proteins specifically in the flowering period may provide protection.

For such genetic engineering approaches to be highly efficient it is advantageous for there to be concentrated expression of the appropriate nucleic acid sequence to be expressed transgenically in particular in the petals of the flower. Constitutive expression in the whole plant may make the effect problematic, for example through dilution, or impair the growth of the plant or the quality of the plant product. In addition, there may through constitutive expression be increased switching-off of the transgene (''gene silencing'').

Promoters having specificity for the flower are advantageous in this connection. The skilled worker is aware of a large number of proteins whose recombinant expression in the flower is advantageous. The skilled worker is also aware of a large number of genes through which advantageous effects can likewise be achieved through repression or switching-off thereof by means of expression of a corresponding antisense RNA. Non-restrictive examples of advantageous effects which may be mentioned are: achieving resistance to abiotic stress factors (heat, cold, aridity, increased moisture, environmental toxins, UV radiation) and biotic stress factors (pathogens, viruses, insects and diseases), improving the properties of human and animal foods, improving the growth rate or the yield, achieving a longer or earlier flowing period, altering or enhancing the scent or the coloring of the flowers. Non-restricive examples of the nucleic acid sequences or polypeptides which can be employed in these applications and which may be mentioned are:

1. Improved UV protection of the flowers of plants through alteration of the pigmentation through expression of particular polypeptides such as enzymes or regulators of flavonoid biosynthesis (e.g. chalcone synthases, phenylalanine ammonia-lyases), of DNA repair (e.g.

photolyases; Sakamoto A et al.(1998) DNA Seq 9(5-6):335-40), of isoprenoid biosynthesis (e.g. deoxyxylulose-5-phosphate synthases), of IPP synthesis or of carotenoid biosynthesis (e.g. phytoene synthases, phytoene desaturases, lycopene cyclases, hydroxylases or ketolases). Preference is given to nucleic acids which code for the Arabidopsis thaliana chalcone synthase (GenBank Acc. No.: M20308), the Arabidopsis thaliana 6-4 photolyase (GenBank Acc. No.:BAB00748) or the Arabidopsis thaliana blue light photoreceptor/photolyase homolog (PHH1) (GenBank Acc. No.: U62549) or functional equivalents thereof.

2. Improved protection of the flower of plants from abiotic stress factors such as aridity, heat or cold, for example through overexpression of the antifreeze polypeptides (e.g. from Myoxocephalus scorpius; WO 00/00512), of the Arabidopsis thaliana transcription activator CBF1, glutamate dehydrogenases (WO 97/12983, WO 98/11240), a late embryogenesis gene (LEA), for example from barley (WO 97/13843), calcium-dependent protein kinase genes (WO 98/26045), calcineurins (WO 99/05902), farnesyl transferases (WO 99/06580; Pei ZM et al. (1998) Science 282:287-290), ferritin (Deak M et al. (1999) Nature Biotechnology 17:192-196), oxalate oxidase (WO 99/04013; Dunwell JM (1998) Biotechnology and Genetic Engeneering Reviews 15:1-32), DREB1A factor (dehydration response element B 1A; Kasuga M et al. (1999) Nature Biotechnology 17:276-286), genes of mannitol or trehalose synthesis (e.g. trehalose-phosphate synthases; trehalose-phosphate phosphatases, WO 97/42326); or through inhibition of genes such as of trehalase (WO 97/50561). Particular preference is given to nucleic acids which code for the Arabidopsis thaliana transcriptional activator CBF1 (Gen-Bank Acc. No.: U77378) or the antifreeze protein from Myoxocephalus octodecemspinosus (GenBank Acc. No.: AF306348) or functional equivalents thereof.

3. Achieving resistance for example to fungi, insects, nematodes and diseases through targeted secretion or accumulation of certain metabolites or proteins in the flower. Examples which may be mentioned are glucosinolates (nematode defense), chitinases or glucanases and other enzymes which destroy the cell wall of parasites, ribosome-inactivating proteins (RIPs) and other proteins of the plant resistance and stress response, like those induced on injury or microbial attack of plants or chemically by, for example, salicylic acid,

jasmonic acid or ethylene, lysozymes from non-plant sources such as, for example, T4 lysozyme or lysozme from various mammals, insecticidal proteins such as Bacillus thuringiensis endotoxin, α-amylase inhibitor or protease inhibitors (cowpea trypsin inhibitor), glucanases, lectins (e.g. phytohemagglutinin, snowdrop lectin, wheatgerm agglutinin), RNAses or ribozymes. Particular preference is given to nucleic acids which code for the chit42 endochitinase from Trichoderma harzianum (GenBank Acc. No.: S78423) or for the N-hydroxylating, multifunctional cytochrome P-450 (CYP79) from Sorghum bicolor (GenBank Acc. No.: U32624) or functional equivalents thereof.

4. Achieving defense against or attraction of insects, for example through increased release of volatile scents or messengers through, for example, enzymes of terpene biosynthesis.

5. Achieving an ability to store in flower tissues which normally contain no storage proteins or lipids, with the aim of increasing the yield of these substances, e.g. by expression of an acetyl-CoA carboxylase or of enzymes for esterification of metabolites. Preference is given to nucleic acids which code for the Medicago sativa acetyl-CoA carboxylase (Accase) (GenBank Acc. No.: L25042) or functional equivalents thereof.

6. Expression of transport proteins which improve the uptake of metabolites, nutrients or water into the flower and thus optimize flower growth, metabolite composition or yield, for example through expression of an amino acid transporter which increases the rate of uptake of amino acids, or of a monosaccharide transporter which promotes the uptake of sugars. Preference is given to nucleic acids which code for the Arabidopsis thaliana cationic amino acid transporter (GenBank Acc. No.: X92657) or for the Arabidopsis thaliana monosaccharide transporter (Gen-Bank Acc. No.: AJ002399) or functional equivalents thereof.

7. Expression of genes which bring about an accumulation of fine chemicals, such as of tocopherols, tocotrienols, phenylpropanoids, isoprenoids or carotenoids, in the flower. Examples which may be mentioned are the deoxyxylulose-5-phosphate synthases, phytoene synthases, lycopene β-cyclases and the β-carotene ketolases. Preference is given to nucleic

acids which code for the Haematococcus pluvialis NIES-144 (Acc. No. D45881) ketolase or functional equivalents thereof.

8. Modification of wax ester formation or of the composition of the deposited oligosaccharides to improve protection against environmental effects or to improve digestibility on use in animal or human foods. An example which may be mentioned is overexpression of endo-xyloglucan transferase. Preference is given to nucleic acids which code for the Arabidopsis thaliana endo-xyloglucan transferase (EXGT-A1) (Gen-Bank Acc. No.:AF163819) or functional equivalents thereof.

9. Expression of genes, DNA binding proteins, dsRNA and antisense constructions for altering the flower morphology, the time of flowering and the flower senescence, and the flower metabolism. Preference is given to constructions which increase the number of petals, e.g. through down regulation of AGAMOUS and its homologous genes (Yanofsky MF et al. (1990) Nature 346:35-39), make the time of flowering earlier, e.g. through down regulation of FLOWERING LOCUS C (FLC) (Tadege M et al. (2001) Plant J 28(5):545-53) or later, e.g. through overexpression of FLC and delay senescence, e.g. through conferring a flower-specific ethylene insensitivity.

10. Generation of sterile plants by preventing pollenation and/or germination by means of the expression of a suitable inhibitor, for example of a toxin, in flowers.

11. Production of nutraceuticals such as, for example,

a) carotenoids and/or phenylpropanoids e.g. through optimization of the flowers' own metabolic pathways, e.g. through expression of enzymes and regulators of isoprenoid biosynthesis. Preference is given to nucleic acids which code for the Arabidopsis thaliana chalcone synthase (GenBank Acc. No.: M20308), the Arabidopsis thaliana 6-4 photolyase (GenBank Acc.No.:BAB00748) or the Arabidopsis thaliana blue light photoreceptor/photolyase homolog (PHH1) (GenBank Acc. No.: U62549) or functional equivalents thereof. Preference is likewise given to nucleic acids which code for enzymes and regulators of isoprenoid biosynthesis such as the deoxyxylulose-5-phosphate synthases and of carotenoid biosynthesis such as the phytoene synthases, lycopene cyclases and ketolases, such as of tocopherols, tocotrienols, phenylpropanoids, isoprenoids or carotenoids, in the

flower. Examples which may be mentioned are the deoxyxylulose-5-phosphate synthases, phytoene synthases, lycopene cyclases and the carotene ketolases. Particular preference is given to nucleic acids which code for the Haematococcus pluvialis, NIES-144 (Acc. No. D45881) ketolase or functional equivalents.

b) polyunsaturated fatty acids such as, for example, arachidonic acid or EPA (eicosapentaenoic acid) or DHA (docosahexaenoic acid) through expression of fatty acid elongases and/or desaturases or production of proteins having improved nutritional value, such as, for example, having a high content of essential amino acids (e.g. the methionine-rich 2S albumin gene of the Brazil nut). Preference is given to nucleic acids which code for the Bertholletia excelsa methionine-rich 2S albumin (GenBank Acc. No.: AB044391), the Physcomitrella patens Δ6-acyl lipid desaturase (GenBank Acc. No.: AJ222980; Girke et al. (1998) Plant J 15:39-48), the Mortierella alpina Δ6-desaturase (Sakura-dani et al 1999 Gene 238:445-453), the Caenorhabditis elegans Δ5-desaturase (Michaelson et al. (1998) FEBS Letters 439:215-218), the Caenorhabditis elegans Δ5-fatty-acid desaturase (des-5) (GenBank Acc. No.: AF078796), the Mortierella alpina Δ5-desaturase (Michaelson et al. J Biol Chem 273:19055-19059), the Caenorhabditis elegans Δ6-elongase (Beaudoin et al. (2000) Proc Natl. Acad. Sci. 97:6421-6426), the Physcomitrella patens Δ6-elongase (Zank et al. (2000,) Biochemical Society Transactions 28:654-657) or functional equivalents thereof.

12. Production of pharmaceuticals such as, for example, antibodies, vaccines, hormones and/or antibiotics as described, for example, in Hood EE & Jilka JM (1999) Curr Opin Biotechnol 10(4):382-6; Ma JK & Vine ND (1999) CurrTop Microbiol Immunol 236:275-92.

Further examples of advantageous genes are mentioned for example in Dunwell JM (2000) Transgenic approaches to crop improvement. J Exp Bot. 51 Spec No:487-96.

A further aspect of the invention relates to the use of the transgenic organisms of the invention described above, and of the cells, cell cultures, parts - such as, for example, in the case of transgenic plant organisms roots, leaves etc. - and transgenic

propagation materials such as seeds or fruits, derived therefrom for producing human or animal foods, pharmaceuticals or fine chemicals.

Preference is further given to a method for the recombinant production of pharmaceuticals or fine chemicals in host organisms, where a host organism is transformed with one of the expression cassettes described above, and this expression cassette comprises one or more structural genes which code for the desired fine chemical, or catalyze the biosynthesis thereof, the transformed host organism is cultivated, and the desired fine chemical is isolated from the cultivation medium. This method can be applied widely to fine chemicals such as enzymes, vitamins, amino acids, sugars, fatty acids, natural and synthetic flavorings, aromatizing substances and colorants. Production of tocopherols and tocotrienols, and carotenoids such as, for example, astaxanthin is particularly preferred. Cultivation of the transformed host organisms and isolation from the host organisms or from the cultivation medium takes place by methods known to the skilled worker. The production of pharmaceuticals such as, for example, antibodies or vaccines is described in Hood EE & Jilka JM (1999) Curr Opin Biotechnol 10 (4)382-6; Ma JK & Vine ND (1999) Curr Top Microbiol Immunol 236:275-92.

A further aspect of the invention relates to the use of the ε-cyclase promoter sequences of the invention (preferably the sequences shown in SEQ ID NO: 1, 7 or 8) for reducing the amount of protein, amount of mRNA and/or activity of an ε-cyclase.

Thus, when an ε-cyclase activity is reduced by comparison with the wild type, the amount of lycopene converted, or the amount of δ-carotene formed, in a particular time by the ε-cyclase protein is reduced by comparison with the wild type.

''Reducing'' or ''reduce'' is to be interpreted broadly in connection with an ε-cyclase or the amount of protein, amount of mRNA and/or activity, and includes the partial or substantially complete inhibition or blocking, based on various cell-biological mechanisms, of the functionality of an ε-cyclase in a plant cell, plant or a part, tissue, organ, cells or seeds derived therefrom.

A reduction for the purposes of the invention also includes a quantitative reduction in an ε-cyclase as far as substantially complete absence of the ε-cyclase (i.e. undetectability of ε-cyclase activity or immunological undetectability of the

ε-cyclase). In this connection, a particular ε-cyclase (or the relevant amount of protein, amount of mRNA and/or activity) in a cell or an organism is reduced preferably by at least 5%, further preferably at least 20%, further preferably at least 50%, further preferably 100%. Reduction means in particular also the complete absence of the ε-cyclase (or of its amount of protein, amount of mRNA and/or activity).

Various strategies for reducing the amount of protein, amount of mRNA and/or activity of the ε-cyclase are included according to the invention. The skilled worker will appreciate that a number of different methods are available for influencing the amount of protein, amount of mRNA and/or activity of an ε-cyclase in the desired way. For example, the reduction can be achieved by introducing at least one double-stranded ribonucleic acid sequence which has at least partial homology with the ε-cyclase promoter sequences of the invention (ε-cyclase promoter dsRNA). An alternative possibility is also to attach expression cassettes ensuring dsRNA expression.

The method of gene regulation by means of double-stranded RNA (''double-stranded RNA interference''; dsRNAi) has been described many times for animal and plant organisms (e.g. Matzke MA et al. (2000) Plant Mol Biol 43:401-415; Fire A. et al (1998) Nature 391:806-811; WO 99/32619; WO 99/53050; WO 00/68374; WO 00/44914; WO 00/44895; WO 00/49035; WO 00/63364). Reference is hereby expressly made to the processes and methods described in the indicated citations. dsRNAi methods are based on the phenomenon of simultaneous introduction of strand and complementary strand of a gene transcript bringing about a highly efficient suppression as the expression of the corresponding gene. The resulting phenotype is very similar to that of a corresponding knock-out mutant (Waterhouse PM et al. (1998) Proc Natl Acad Sci USA 95:13959-64).

''Double-stranded RNA molecule'' means for the purposes of the invention preferably one or more ribonucleic acid sequences which are able because of complementary sequences theoretically (e.g. according to the base-pair rules of Watson and Crick) and/or actually (e.g. on the basis of hybridization experiments in vitro and/or in vivo) to form double-stranded RNA structures. The skilled worker is aware that the formation of double-stranded RNA structures represents a dynamic equilibrium. The ratio of double-stranded molecules to corresponding dissociated forms is

preferably at least 1 to 10, preferably 1:1, particularly preferably 5:1, most preferably 10:1.

A further aspect of the invention therefore relates to double-stranded RNA molecules (dsRNA molecules) which, on introduction into a plant organism (or a cell, tissue, organ or propagation material derived therefrom), bring about the reduction of at least one ε-cyclase. The double-stranded RNA molecule for reducing the expression of an ε-cyclase (ε-cyclase dsRNA) in this case preferably includes

a) a sense RNA strand including at least one ribonucleotide sequence which is substantially identical to at least part of a nucleic acid sequence coding for the promoter region of an ε-cyclase, and

b) an antisense RNA strand which is substantially - preferably completely - complementary to the RNA sense strand under a).

The promoter region of the ε-cyclase is preferably described by a sequence as shown in SEQ ID NO: 1, 7 or 8.

''Substantially identical'' means that the dsRNA sequence may also have insertions, deletions and single point mutations compared with the ε-cyclase promoter target sequence, and nevertheless brings about an efficient reduction of expression. The homology (as defined hereinafter) is preferably at least 75%, preferably at least 80%, very particularly preferably at least 90%, most preferably 100%, between the sense strand of an inhibitory dsRNA and at least part of the nucleic acid sequence coding for an ε-cyclase promoter (or between the antisense strand and the complementary strand of a nucleic acid sequence coding for an ε-cyclase promoter). The skilled worker moreover is aware that, in a comparison of homology between RNA and DNA, the bases uracil and thymine are to be regarded as equivalent.

A 100% sequence identity between dsRNA and an ε-cyclase promoter is not absolutely necessary for bringing about an efficient reduction of ε-cyclase expression. Accordingly, there is the advantage that the method is tolerant to sequence differences like those which may be present owing to genetic mutations, polymorphisms or evolutionary divergences.

The length of the partial segment is at least 10 bases, preferably at least 25 bases, particularly preferably at least 50

bases, very particularly preferably at least 100 bases, most preferably at least 200 bases or at least 300 bases.

It is alternatively possible for a ''substantially identical'' dsRNA also to be defined as nucleic acid sequence which is able to hybridize with part of an ε-cyclase gene or promoter sequence (e.g. in 400 mM NaCl, 40 ∝M PIPES pH 6.4, 1 ∝M EDTA at 50°C or 70°C for 12 to 16 h).

''Substantially complementary'' means that the antisense RNA strand may also have insertions, deletions and single point mutations by comparison with the complement of the sense RNA strand. The homology is preferably at least 80%, preferably at least 90%, very particularly preferably at least 95%, most preferably 100%, between the antisense RNA strand and the complement of the sense RNA strand.

''Part of a nucleic acid sequence coding for an ε-cyclase promoter'' means fragments of a nucleic acid sequence coding for an ε-cyclase promoter, preferably the promoter sequences as shown in SEQ ID NO: 1, 2 or 3 or functional equivalents thereof. In this connection, the fragments preferably have a sequence length of at least 20 bases, preferably at least 50 bases, particularly preferably at least 100 bases, very particularly preferably at least 200 bases, most preferably at least 500 bases.

It is particularly advantageous to use the ε-cyclase promoter region to reduce the ε-cyclase activity because only low homologies with other genes are present here, and thus the reduction can be highly specific without effecting the expression of other genes.

The dsRNA can consist of one or more strands of polyribonucleotides. It is, of course, also possible to achieve the same purpose by introducing a plurality of individual dsRNA molecules, each of which include one of the ribonucleotide sequence segments defined above, into the cell or the organism. The double-stranded dsRNA structure can be formed starting from two complementary separate RNA strands or - preferably - starting from a single self-complementary RNA strand. In this case, sense RNA strand and antisense RNA strand are preferably connected together covalently in the form of an inverted repeat.

In a preferred embodiment, a further aspect of the invention includes ribonucleic acid molecules including

a) at least one ribonucleotide sequence which is substantially identical to at least one part of a nucleic acid sequence coding for the promoter region of an ε-cyclase, and

b) at least one further ribonucleotide sequence which is substantially complementary to at least one part of the ribonucleotide sequence under a),

where a) and b) are connected together covalently, and further functional elements may be located where appropriate between a) and b).

The promoter region of the ε-cyclase is preferably described by a sequence as shown in SEQ ID NO: 1, 7 or 8.

As described, for example, in WO 99/53050, the dsRNA may also include a hairpin structure through connection of sense and antisense strands by a connecting sequence (''linker''; for example an intron). The self-complementary dsRNA structures are preferred, because they require merely the expression of one RNA sequence and include the complementary RNA strands always in an equimolar ratio. The connecting sequence is preferably an intron (e.g. an intron of the potato ST-LS1 gene; Vancanneyt GF et al. (1990) Mol Gen Genet 220(2):245-250).

If the two strands of the dsRNA are to be put together in a cell or plant, this can take place in the following way, for example:

a) transformation of the cell or plant with a vector which includes both expression cassettes,

b) cotransformation of the cell or plant with two vectors, where one includes the expression cassettes with the sense strand the other includes the expression cassettes with the antisense strand,

c) crossing of two individual plant lines, where one includes the expression cassettes with the sense strand and the other includes the expression cassettes with the antisense strand.

Formation of the RNA duplex can be initiated either outside the cell or inside it.

The dsRNA can be synthesized either in vivo or in vitro. For this purpose it is possible to put a DNA sequence coding for a dsRNA into an expression cassette under the control of at least one

genetic control element (such as, for example, a promoter). Polyadenylation is unnecessary, nor need any elements be present to initiate translation. The expression cassette for the ε-cyclase promoter dsRNA is preferably contained on the expression vector. The invention includes corresponding expression vectors.

In a particular preferred embodiment, expression of the dsRNA takes place starting from an expression construct under the functional control of a flower-specific promoter. The promoter employed in this connection is preferably not the ε-cyclase promoter from which the dsRNA has been derived. However, it is very possible for it to be an ε-cyclase promoter of a different species. Thus, for example, the sunflower ε-cyclase promoter could be used to express the dsRNA derived from the Tagetes erecta ε-cyclase promoter. However, expression of the dsRNA derived from an ε-cyclase promoter is preferably under the control of a promoter which is not an ε-cyclase promoter, particularly preferably under the control of the Cucumis sativus CHRC promoter (SEQ ID NO: 81) or of the AP3P promoter (SEQ ID NO: 77) or of a functionally equivalent part thereof.

The expression cassettes coding for the antisense and/or the sense strand of an ε-cyclase dsRNA or for the self-complementary strand of the dsRNA are for this purposes preferably inserted into a transformation vector and introduced into the plant cell using the methods described below. Stable insertion into the genome is advantageous for the method of the invention.

The dsRNA can be introduced in an amount which makes at least one copy possible per cell. Larger amounts (e.g. at least 5, 10, 100, 500 or 1000 copies per cell) may where appropriate bring about a more efficient reduction.

The invention also includes methods for producing ketocarotenoids, where the amount of mRNA and/or activity of at least one ε-cyclase is reduced by introducing at least one of the double-stranded RNA sequences or ribonucleic acid sequences of the invention or an expression cassette or expression cassettes ensuring expression thereof.

Ketocarotenoids means carotenoids which comprise at least one keto group, such as, for example, astaxanthin, canthaxanthin, echinenone, 3-hydroxyechinenone, 3'-hydroxyechinenone, adonirubin and adonixanthin.

Sequences

1. SEQ ID NO: 1 nucleic acid sequence coding for the Tagetes erecta ε-cyclase promoter

2. SEQ ID NO: 2 nucleic acid sequence coding for the Tagetes erecta ε-cyclase promoter including the ε-cyclase 5'-untranslated region

3. SEQ ID NO: 3 nucleic acid sequence coding for the Tagetes erecta ε-cyclase promoter including 5'-untranslated region and region coding for the transit peptide

4. SEQ ID NO: 4 amino acid sequence of the putative Tagetes erecta ε-cyclase transit peptide

5. SEQ ID NO: 5 nucleic acid sequence coding for the Tagetes erecta ε-cyclase promoter including the ε-cyclase 5'-untranslated region flanked by restriction cleavage sites for cloning

6. SEQ ID NO: 6 nucleic acid sequence coding for the Tagetes erecta ε-cyclase promoter including 5'-untranslated region and region coding for the transit peptide flanked by restriction cleavage sites for cloning

7. SEQ ID NO: 7 nucleic acid sequence coding for the Arabidopsis thaliana ε-cyclase promoter including the ε-cyclase 5'-untranslated region

8. SEQ ID NO: 8 nucleic acid sequence coding for the Oryza sativa ε-cyclase promoter including the ε-cyclase 5'-untranslated region

9. SEQ ID NO: 9 nucleic acid sequence coding for a Tagetes erecta ε-cyclase

10. SEQ ID NO: 10 amino acid sequence of a Tagetes erecta ε-cyclase

11. SEQ ID NO: 11 nucleic acid sequence coding for a Tagetes erecta ε-cyclase

12. SEQ ID NO: 12 amino acid sequence of a Tagetes erecta ε-cyclase

13. SEQ ID NO: 13 nucleic acid sequence coding for an Arabidopsis thaliana ε-cyclase

14. SEQ ID NO: 14 amino acid sequence of an Arabidopsis thaliana ε-cyclase

15. SEQ ID NO: 15 nucleic acid sequence coding for a rice ε-cyclase

16. SEQ ID NO: 16 amino acid sequence of a rice ε-cyclase

17.-22 SEQ ID NO: 17 to 22: sequence motifs for ε-cyclase proteins

23. SEQ ID NO: 23 nucleic acid sequence coding for a Lactuca sativa ε-cyclase (homologous sequence H1)

24. SEQ ID NO: 24 amino acid sequence of a Lactuca sativa ε-cyclase (homologous sequence H1)

25. SEQ ID NO: 25 nucleic acid sequence coding for an Adonis palaestina ε-cyclase (homologous sequence H2)

26. SEQ ID NO: 26 amino acid sequence of an Adonis palaestina ε-cyclase (homologous sequence H2)

27. SEQ ID NO: 27 nucleic acid sequence coding for an Adonis palaestina ε-cyclase (homologous sequence H3)

28. SEQ ID NO: 28 amino acid sequence of an Adonis palaestina ε-cyclase (homologous sequence H3)

29. SEQ ID NO: 29 nucleic acid sequence coding for an Arabidopsis thaliana ε-cyclase (homologous sequence H4)

30. SEQ ID NO: 30 amino acid sequence of an Arabidopsis thaliana ε-cyclase (homologous sequence H4)

31. SEQ ID NO: 31 nucleic acid sequence coding for a Citrus X paradisi ε-cyclase (homologous sequence H5)

32. SEQ ID NO: 32 amino acid sequence of Citrus X paradisi ε-cyclase (homologous sequence H5)

33. SEQ ID NO: 33 nucleic acid sequence coding for a Citrus X paradisi ε-cyclase (homologous sequence H6)

34. SEQ ID NO: 34 amino acid sequence of Citrus X paradisi ε-cyclase (homologous sequence H6)

35. SEQ ID NO: 35 nucleic acid sequence coding for a Citrus sinensis ε-cyclase (homologous sequence H7)

36. SEQ ID NO: 36 amino acid sequence of a Citrus sinensis ε-cyclase (homologous sequence H7)

37. SEQ ID NO: 37 nucleic acid sequence coding for a Spinacea oleracea ε-cyclase (homologous sequence H8)

38. SEQ ID NO: 38 amino acid sequence of a Spinacea oleracea ε-cyclase (homologous sequence H8)

39. SEQ ID NO: 39 nucleic acid sequence coding for a Solanum tuberosum ε-cyclase (homologous sequence H9)

40. SEQ ID NO: 40 amino acid sequence of a Solanum tuberosum ε-cyclase (homologous sequence H9)

41. SEQ ID NO: 41 nucleic acid sequence coding for a Daucus carota ε-cyclase (homologous sequence H10)

42. SEQ ID NO: 42 amino acid sequence of a Daucus carota ε-cyclase (homologous sequence H10)

43. SEQ ID NO: 43 nucleic acid sequence coding for a Daucus carota ε-cyclase (homologous sequence H11)

44. SEQ ID NO: 44 amino acid sequence of a Daucus carota ε-cyclase (homologous sequence H11)

45. SEQ ID NO: 45 nucleic acid sequence coding for a tomato ε-cyclase (homologous sequence H12)

46. SEQ ID NO: 46 amino acid sequence of a tomato ε-cyclase (homologous sequence H12)

47. SEQ ID NO: 47 nucleic acid sequence coding for ε-cyclase-specific probe (gecyc1; 510 bp)

48. SEQ ID NO: 48 oligonucleotide primer PR16
5'-ggcacgaggcaaagcaaagg-3'

49. SEQ ID NO: 49 oligonucleotide primer PR22
5'-cgataagtgcgacattcaagc-3'

50. SEQ ID NO: 50 nucleic acid sequence including part of the Tagetes erecta ε-cyclase promoter obtained by iPCR

51. SEQ ID NO: 51 nucleic acid sequence including part of the Tagetes erecta ε-cyclase promoter obtained by TAIL PCR

52. SEQ ID NO: 52 oligonucleotide primer PR50
5'-cgccttgtatctgtttggattgg-3'

53. SEQ ID NO: 53 oligonucleotide primer PR51
5'-ctaacaatcaatgagtatgagagc-3'

54. SEQ ID NO: 54 oligonucleotide primer PR60
5'-agagcaaggccagcaggaccacaacc-3'

55. SEQ ID NO: 55 oligonucleotide primer PR61
5'-ccttgggagcttttgggataggctag-3'

56. SEQ ID NO: 56 oligonucleotide primer PR63
5'-tcacgccttgtatctgtttggattgg-3'

57. SEQ ID NO: 57 oligonucleotide primer from the set of AD1 primers as was found in the amplicon 5'-gtcgagtatggagtt-3'

58. SEQ ID NO: 58 nucleic acid sequence encoding iPCR fragment (734 bp) from pTA-ecycP

59. SEQ ID NO: 59 oligonucleotide primer OL1
5'-ctcgagagtaaaatcgttagttatg-3'

60. SEQ ID NO: 60 oligonucleotide primer OL2
5'-ccatggccattgattgttagtaatgattc-3'

61. SEQ ID NO: 61 oligonucleotide primer OL3
5'-ccatggtaatttgcttcgtgtatctgatg-3'

62. SEQ ID NO: 62 oligonucleotide primer OL4
5'-ccatggcgctagcagcgacagtaatg-3'

63. SEQ ID NO: 63 oligonucleotide primer OL5
5'-gatatccggtgtgagggaactag-3'

64. SEQ ID NO: 64 oligonucleotide primer PR1
5'-gcaagctcgacagctacaaacc-3'

65. SEQ ID NO: 65 oligonucleotide primer PR2
5'-gaagcatgcagctagcagcgacag-3'

66. SEQ ID NO: 66 nucleic acid sequence coding for Ketolase-35S terminator construct

67. SEQ ID NO: 67 oligonucleotide primer PR7
5'-gagctcactc actgatttcc attgcttg-3'

68. SEQ ID NO: 68 oligonucleotide primer PR8
5'-cgccgttaagtcgatgtccgttgatttaaacagtgtc-3'

69. SEQ ID NO: 69 oligonucleotide primer PR9
5'-atcaacggac atcgacttaa cggcgtttgt aaac-3'

70. SEQ ID NO: 70 oligonucleotide primer PR10
5'-taagcttttt gttgaagaga tttgg-3'

71. SEQ ID NO: 71 oligonucleotide primer PR40
5'-gtcgactacg taagtttctg cttctacc-3'

72. SEQ ID NO: 72 oligonucleotide primer PR41
5'-ggatccggtg atacctgcac atcaac-3'

73. SEQ ID NO: 73 oligonucleotide primer PR124
5'-aagcttaccg atagtaaaat cgttagtt-3'

74. SEQ ID NO: 74 oligonucleotide primer PR125
5'-ctcgagctta ccgatagtaa aatcgttagt t-3'

75. SEQ ID NO: 75 oligonucleotide primer PR126
5'-gtcgacaaca acaacaaaca acctttgc-3'

76. SEQ ID NO: 76 oligonucleotide primer PR127
5'-ggatccaaca acaacaaaca acctttgc-3'

77. SEQ ID NO: 77 nucleic acid sequence coding for a modified version (AP3P) of the flower-specific Arabidopsis thaliana promoter AP3

78. SEQ ID NO: 78 nucleic acid sequence coding for PIV2 intron of the potato ST-LS1 gene.

79. SEQ ID NO: 79 nucleic acid sequence coding for the sense strand of the dsRNA directed against the ε-cyclase promoter

80. SEQ ID NO: 80 nucleic acid sequence coding for the antisense strand of the dsRNA directed against the ε-cyclase promoter

81. SEQ ID NO: 81 nucleic acid sequence coding for the Cucumis sativus chromoplast-specific carotenoid-associated protein (CHRC) promoter

82. SEQ ID NO: 82 oligonucleotide primer PRCHRC5
5'-gagctctaca aattagggtt ac-3'

83. SEQ ID NO: 83 oligonucleotide primer PRCHRC3
5'-aagcttatta tttccaaatt ccg-3'

Figures

The general abbreviations used in the following figures have the following meaning:

GUSI-Intron-GUSII: reporter gene (bacterial β-glucuronidase)
Intron: Intron
NosT: nopaline synthase (NOS) terminator sequence
RB/LB: right or left T-DNA border
35-T: 35S CaMV terminator
NptII: canamycin resistence
NosP: nopaline synthase (NOS) promoter sequence
aadA: bacterial spectinomycin resistance
colE1: origin of replication

1. Fig. 1: analysis of the ε-cyclase transcript level total RNA isolated from leaves (L) and flower stages (1-7) of Tagetes erecta by means of RNA gel blotting analysis

2. Fig. 2: diagrammatic representation of the pEcycP1:GUS vector for flower-specific expression of the β-glucuronidase reporter gene (GUS) under the control of the Tagetes erecta ecycP1 regulatory element (promoter and 5'-untranslated region)

ecycP1: Tagetes erecta ε-cyclase promoter including 5'-untranslated region (SEQ ID NO: 2)

3. Fig.3: diagrammatic representation of the pEcycP2:GUS vector for flower-specific expression of the β-glucuronidase reporter gene (GUS) under the control of the Tagetes erecta ecycP2 regulatory element (promoter and 5'-untranslated region and transit peptide)

ecycP2: Tagetes erecta ε-cyclase promoter including 5'-untranslated region and transit peptide (SEQ ID NO: 3)

4. Fig.4: diagrammatic representation of the pEcycP2:KETO vector for flower-specific expression of the Haematococcus pluvialis ketolase (KETO; SEQ ID NO: 66) under the control of the Tagetes erecta ecycP2 regulatory element (promoter and 5'-untranslated region and transit peptide; SEQ ID NO: 3).

5. Fig.5: diagrammatic representation of the pS5AI7 vector for flower-specific expression of ε-cyclase promoter specific dsRNA under the control of the AP3P promoter fragment for flower-specific reduction of the ε-cyclase transcript level.
AP3P: modified AP3P promoter (777 bp),
P-sense: 358 bp ε-cyclase promoter fragment in sense orientation,
intron: IV2 intron of the potato ST-LS1 gene
P-anti: the 361 bp ε-cyclase promoter fragment in antisense orientation.

6. Fig.6: diagrammatic representation of the pS5CI7 vector for flower-specific expression of ε-cyclase promoter specific dsRNA under the control of the CHRC promoter fragment for flower-specific reduction of the ε-cyclase transcript level

CHRC: CHRC promoter (1537 bp),
P-sense: 358 bp ε-cyclase promoter fragment in sense orientation,

intron: IV2 intron of the potato ST-LS1 gene
P-anti: the 361 bp ε-cyclase promoter fragment in antisense orientation.

7. Fig.7: iPCR amplicon comprising the 312 bp fragment of the ε-cyclase promoter

8. Fig.8: TAIL PCR amplicon comprising the 199 bp fragment of the ε-cyclase promoter

9. Fig.9: nucleotide sequence comparison between the published sequence of the Haematococcus pluvialis ketolase (GenBank Acc. No.: X86782) and the sequence provided within the scope of the invention (cf. example 3).

10. Fig.10: protein sequence comparison between the published sequence of the Haematococcus pluvialis ketolase (GenBank Acc. No.: X86782) and the sequence provided within the scope of the invention (cf. example 3).

11. Fig.11: cloning cassette for producing inverted repeat expression cassettes for flower-specific expression of ε-cyclase dsRNAs.
AP3P: modified AP3P promoter (777 bp),
rbcs: pea rbcS transit peptide (206 bp),
intron: PIV2 intron of the ST-LS1 gene (SEQ ID NO: 78)
term: CaMV 35S polyadenylation signal (762 bp).

12. Fig.12A-C: sequence comparison of various plant ε-cyclases.
A: GenBank Acc. No.: AF152246 (524) Citrus x paradisi ''lycopene cyclase''
B: GenBank Acc. No.: AF212130 (165) Daucus carota partial ecyclase sequence
C: GenBank Acc. No.: AF229684 (201) Daucus carota partial ecyclase sequence
D: GenBank Acc. No.: AF251016 (516) Tagetes erecta ecyclase
E: GenBank Acc. No.: AF321535 (529) Adonis palaestina ecyclase
F: GenBank Acc. No.: AF321536 (529) Adonis palaestina ecyclase
G: GenBank Acc. No.: AF321537 (382) Solanum tuberosum partial ecyclase sequence
H: GenBank Acc. No.: AF321538 (533) Lactuca sativa ecyclase
I: GenBank Acc. No.: AF450280 (262) Citrus sinensis ecyclase

J: GenBank Acc. No.: AF463497 (517) Spinacea oleracea ecyclase
K: GenBank Acc. No.: AF486650 (437) Citrus x paradisi ecyclase
L: GenBank Acc. No.: AP003332 (540) rice ecyclase
M: GenBank Acc. No.: AY099485 (525) Tagetes erecta ecyclase
N: GenBank Acc. No.: L40176 (501) Arabidopsis ''lycopene cyclase''
O: GenBank Acc. No.: NM125085 (524) Arabidopsis ecyclase
P: GenBank Acc. No.: O65837 ecyclase (526) tomato

13. Fig.13: diagrammatic representation of the inverse PCR (''iPCR'')
For the ''iPCR'', genomic DNA of a target organism having the promoter sequence to be isolated is completely digested with a given restriction enzyme, and then the individual fragments are religated, i.e. connected together to form a circular molecule, in a diluted mixture. The large number of resulting circular DNA molecules includes those comprising the known sequence (i.e. the sequence coding for a homologous protein). The circular molecule can be amplified, starting therefrom, by means of PCR using a primer pair in which both primers are able to anneal to the known sequence segment.
Abbreviations: P - promoter sequence; CR - coding region; L - ligation site; PCR - polymerase chain reaction. Arrows represent the binding site of potential oligonucleotide primers in the area of the coding region.

Examples

General methods:

Oligonucleotides can be chemically synthesized for example in a known manner by the phosphoramidite method (Voet & Voet (1995), 2nd edition, Wiley Press New York, pages 896-897). The cloning steps carried out for the purposes of the present invention, such as, for example, restriction cleavages, agarose gel electrophoresis, purification of DNA fragments, transfer of nucleic acids to nitrocellulose and nylon membranes, linkage of DNA fragments, transformation of E. coli cells, culturing of bacteria, replication of phages and sequence analysis of recombinant DNA, are carried out as described in Sambrook et al. (1989) Cold Spring Harbor Laboratory Press; ISBN 0-87969-309-6. Recombinant DNA molecules are sequenced by the method of Sanger (Sanger et al.(1977) Pro Natl Acad Sci USA 74:5463-5467) using an ABI laser fluorescence DNA sequencer.

Example 1: Analysis of ε-cyclase RNA transcript levels during the development of Tagetes erecta flowers

Total RNA from Tagetes erecta leaves and flowers is prepared by harvesting plant tissue, freezing it in liquid nitrogen and powdering in a mortar. 100 mg of the frozen, powdered plant tissue are then transferred into a reaction vessel and taken up in 0.8 ml of Trizol® buffer (LifeTechnologies). The suspension is extracted with 0.2 ml of chloroform. After centrifugation at 12 000 g for 15 minutes, the aqueous supernatant is removed and transferred into a new reaction vessel and extracted with one volume of ethanol. The RNA is precipitated with one volume of isopropanol and washed with 75% ethanol, and the pellet is dissolved in DEPC water (water incubated with 1/1000 volume of diethyl pyrocarbonate (DEPC) at room temperature overnight and then autoclaved). The RNA concentration is determined by photometry.

The relative amount of ε-cyclase transcript in Tagetes leaves and flower stages is analyzed by RNA gel blotting as described in Sambrook & Russel (2001, Molecular Cloning: A laboratory manual, 3rd Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, Chapter 7, Protocol 6): about 10 to 15 αg of total RNA of each sample are fractionated in a formaldehyde agarose gel. The relative amounts of total RNA can be estimated from the rRNA bands stained with ethidium bromide (Fig. 1A). The amounts of ε-cyclase transcript are estimated by transferring the fractionated RNA by capillary blotting to a nylon membrane.

A radiolabeled ε-cyclase-specific probe was prepared by amplifying the fragment of SEQ ID NO: 47 (gecyc1) by polymerase chain reaction (PCR) from Tagetes erecta genomic DNA using a sense-specific primer (PR16 = 5'-ggcacgaggcaaagcaaagg-3', SEQ ID NO: 48) and an antisense-specific primer (PR22 = 5'-cgataagtgcgacattcaagc-3', SEQ ID NO: 49).

Tagetes erecta genomic DNA is prepared by harvesting leaf material from Tagetes erecta, freezing it in liquid nitrogen and powdering in a mortar. 100 mg of the frozen, powdered plant tissue is then transferred into a reaction vessel, taken up in 0.75 ml of extraction buffer and incubated at 65_C for 60 min. The extraction buffer is freshly prepared from 25 ml of buffer 1 (0.35 M sorbitol, 0.1 M tris base, 5 αM EDTA, pH 7.5), 25 ml of buffer 2 (0.2 M tris base, 0.05 M EDTA, 2 M NaCL, 2% CTAB), 10 ml of 5% N-lauroylsarcosine sodium) and 0.24 g of sodium bisulfite.

Incubation at 65_C is followed by mixing the suspension with 0.7 ml of chloroform/isoamyl alcohol (24:1) and then centrifuging at 10 000 g for 5 min. The upper aqueous phase is transferred into a new reaction vessel, and the chloroform/isoamyl alcohol extraction is repeated as described. The upper aqueous phase is then transferred into a new reaction vessel, and the DNA is pelleted by adding 1 ml of isopropanol and then centrifuging at 10 000 g for 5 min. The DNA pellet is washed with 0.5 ml of 75% ethanol, then dried and subsequently resuspended in 0.05 ml of sterile water by incubation at 65_C for 5 minutes.

The PCR conditions for amplification of an ε-cyclase-specific fragment from Tagetes erecta genomic DNA are as follows:

The PCR for amplifying an ε-cyclase-specific fragment takes place in a 50 µl reaction mixture which contains:

- 1 µg Tagetes erecta genomic DNA
- 0.25 µM dNTPs
- 0.2 µM primer PR16 (SEQ ID NO: 48)
- 0.2 µM primer PR22 (SEQ ID NO: 49)
- 5 µl 10X PCR buffer (TAKARA)
- 0.25 µl R Taq polymerase (TAKARA)
- 25.8 µl sterile distilled water

The PCR was carried out under the following cycle conditions: 1 cycle at 94_C for 2 minutes. 35 cycles at 94_C for 1 minute, 51_C for 2 minutes and 72_C for 3 minutes. Finally one cycle at 72_C for 10 minutes.

The PCR amplification with PR16 and PR22 results in a 510 bp fragment (SEQ ID NO: 47) which, under stringent hybridization conditions, hybridizes specifically with the ε-cyclase but not with the lycopene β-cyclase from Tagetes erecta. The amplification product is purified using the NucleoSpin® extract kit (Machery & Nagel) as stated by the manufacturer and employed for a radiolabeling reaction with the Highprime® kit (Boehringer Mannheim) as stated by the manufacturer. The prehybridization, hybridization and washing steps are carried out as described in Sambrook & Russel (2001, Molecular Cloning: A laboratory manual, 3rd Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, Chapter 6, Protocol 10). The last washing step with 0.1 x SSC/0.1% SDS at 65_C makes the hybridization highly stringent, sufficient for specific detection with the probe described of ε-cyclase but not lycopene β-cyclase. The relative

ε-cyclase transcript levels can be estimated from the hybridization signals detected with the aid of a phosphoimager. As is evident in Fig. 1B, under the given experimental conditions, the ε-cyclase transcript levels in the leaves are below the limit of detection, whereas large amounts of ε-cyclase transcripts are detectable throughout flower development.

Example 2: Cloning of the ε-cyclase promoter

A 199 bp fragment or the 312 bp fragment of the Tagetes erecta ε-cyclase promoter can be isolated by two independent cloning strategies, inverse PCR (iPCR; adapted from Long et al. Proc Natl Acad Sci USA 90: 10370) and TAIL PCR (Liu YG et al. (1995) Plant J 8: 457-463) using genomic DNA (as described above) from the *Tagetes erecta* line Orangenprinz.

For the iPCR mixture, 2 µg of genomic DNA are digested in a 25 µl reaction mixture with EcoRV and RsaI, then diluted to 300 µl and religated with 3U of ligase at 16_C overnight. PCR amplification using the primers PR50 (SEQ ID NO: 52) and PR51 (SEQ ID NO: 53) produces a fragment which comprises, in each case in sense orientation, 354 bp of the ε-cyclase cDNA (Genbank Acc. No.: AF251016) ligated to 312 bp of the ε-cyclase promoter and 70 bp of the 5'-terminal region of the ε-cyclase cDNA (see Fig. 7).

The conditions for the PCR reactions are as follows:

The PCR for amplifying the PR50-PR51 DNA fragment which comprises inter alia the 312 bp ε-cyclase promoter fragment takes place in a 50 µl reaction mixture containing:

- 1 µl ligation mixture (prepared as described above)
- 0.25 µM dNTPs
- 0.2 µM primer PR50 (SEQ ID NO: 52)
- 0.2 µM primer PR51 (SEQ ID NO: 53)
- 5 µl 10X PCR buffer (TAKARA)
- 0.25 µl R Taq polymerase (TAKARA)
- 28.8 µl sterile distilled water

The PCR reactions are carried out under the following cycle conditions: 1 cycle at 94_C for 2 minutes. 35 cycles at 94_C for 1 minute, 53_C for 1 minute and 72_C for 1 minute. Finally 1 cycle at 72_C for 10 minutes.

The PCR amplification with primer PR50 and PR51 results in a 734 bp fragment which comprises inter alia the 312 bp ε-cyclase promoter fragment (Fig. 7). The amplicon is cloned using standard methods into the PCR cloning vector pCR2.1 (Invitrogen). Sequencings using the primers M13 and T7 result in the sequence SEQ ID NO: 50 for the amplicon.

For the TAIL PCR approach, three successive PCR reactions are carried out each with different gene-specific primers (''nested primers'').

The TAIL1 PCR takes place in a 20 ∝l reaction mixture containing:

- 100 ng genomic DNA (prepared as described above)
- 0.2 ∝M each dNTP
- 0.2 ∝M primer PR60 (SEQ ID NO: 54)
- 0.2 ∝M AD1 primer mixture
- 2 ∝l 10X PCR buffer (TAKARA)
- 0.5 U R Taq polymerase (TAKARA)
- made up to 20 ∝l with sterile distilled water

The AD1 primer mixture initially represented a mixture of primers of the sequences

5'-(a/c/g/t)tcga(g/c)t(a/t)t(g/c)g(a/t)gtt-3'. The primer with the SEQ ID NO: 57 was found in the resulting amplicon.

The TAIL1 PCR reaction was carried out under the following cycle conditions:

- 1 cycle at 93_C for 1 minute and 95_C for 1 minute,
- 5 cycles at 94_C for 30 seconds, 62_C for 1 minute and 72_C for 2.5 minutes,
- 1 cycle at 94_C for 30 seconds, 25_C for 3 minutes, then a temperature increase to 72_C over the course of 3 minutes, 72_C for 2.5 minutes
- 15 cycles at 94_C for 10 seconds, 68_C for 1 minute and 72_C for 2.5 minutes; 94_C for 10 seconds, 68_C for 1 minute and 72_C for 2.5 minutes; 94_C for 10 seconds, 29_C for 1 minute and 72_C for 2.5 minutes;
- 1 cycle at 72_ for 5 minutes.

The TAIL2 PCR takes place in a 21 ∝l reaction mixture containing;

- 1 ∝l of a 1:50 dilution of the TAIL1 reaction mixture (prepared as described above)

- 0.8 ∝M dNTP
- 0.2 ∝M primer PR61 (SEQ ID NO: 55)
- 0.2 ∝M primer AD1 (SEQ ID NO: 57)
- 2 ∝l 10X PCR buffer (TAKARA)
- 0.5 U R Taq polymerase (TAKARA)
- made up to 21 ∝l with sterile distilled water

The TAIL2 PCR reaction is carried out under the following cycle conditions:

- 12 cycles at 94_C for 10 seconds, 64_C for 1 minute, 72_C for 2.5 minutes, 94_C for 10 seconds, 64_C for 1 minute, 72_C for 2.5 minutes; 94_C for 10 seconds, 29_C for 1 minute, 72_C for 2.5 minutes;
- 1 cycle at 72_C for 5 minutes.

The TAIL3 PCR takes place a 100 ∝l reaction mixture containing:

- 1 ∝l of a 1:10 dilution of the TAIL2 reaction mixture (prepared as described above)
- 0.8 ∝M dNTP
- 0.2 ∝M primer PR63 (SEQ ID NO: 56)
- 0.2 ∝M primer AD1 (SEQ ID NO: 57)
- 10 ∝l 10X PCR buffer (TAKARA)
- 0.5 U R Taq polymerase (TAKARA)
- made up to 100 ∝l with sterile distilled water

The TAIL3 PCR reaction is carried out under the following cycle conditions:

- 20 cycles at 94_C for 15 seconds, 29_C for 30 seconds, 72_C for 2 minutes;
- 1 cycle at 72_C for 5 minutes.

The PCR amplification with primer PR63 and AD1 results in a 280 bp fragment which comprises inter alia the 199 bp ε-cyclase promoter fragment (Fig. 8).

The amplicon was cloned using standard methods into the PCR cloning vector pCR2.1 (Invitrogen). Sequencings using the primers M13 and T7 result in the sequence SEQ ID NO: 51. This sequence is identical in the overlap region to the sequence of SEQ ID NO: 50 which is isolated using the iPCR strategy, and thus represents the nucleotide sequence in the Tagetes erecta line Orangenprinz used.

The pCR2.1 clone which contains the 734 bp fragment (SEQ ID NO: 58) which is isolated by the iPCR strategy is called pTA-ecycP and is used to produce the expression constructs.

Example 3: Production of transgenic ε-cyclase expression cassettes and expression vectors

The ε-cyclase regulatory element ecycP1 containing a promoter fragment and the 5'-untranslated region of the Tagetes erecta ε-cyclase is used to express β-glucuronidase (Jefferson et al. (1987) EMBO J 6:3901-3907) in tomato flowers (*Lycopersicon esculentum*). In addition, the ε-cyclase regulatory element ecycP2 containing a promoter fragment, the 5'-untranslated region and the putative transit peptide of the Tagetes erecta ε-cyclase is used to express either β-glucuronidase or the Haematococcus pluvialis ketolase in plastids of tomato flowers.

The transgenic expression vectors pEcycP1:GUS, pEcycP2:GUS, pEcycP2:KETO for the agrobacterium-mediated transformation into *Lycopersicon esculentum* were produced using the binary vector pS0301 (WO 02/00900). The transformation plasmids are produced by producing the fragments ecycP1 and ecycP2 by PCR using the clone pTA-ecycP and the primers OL1 (SEQ ID NO: 59) and OL2 (SEQ ID NO: 60) (for ecycP1) or the primers OL1 (SEQ ID NO: 59) and OL3 (SEQ ID NO: 61) (for ecycP2).

The PCR for amplifying an ε-cyclase-specific fragment takes place in a 50 µl reaction mixture containing:

- 50 ng pTA-ecycP plasmid
- 0.25 µM dNTPs
- 0.2 µM primer OL1 (SEQ ID NO: 59)
- 0.2 µM primer OL2 (SEQ ID NO: 60) for ecycP1 or primer OL3 (SEQ ID NO: 61) for ecycP2
- 5 µl 10X PCR buffer (TAKARA)
- 0.25 µl R Taq polymerase (TAKARA)
- 25.8 µl sterile distilled water.

The PCR is carried out under the following cycle conditions: 1 cycle at 94_C for 2 minutes, 35 cycles at 94_C for 1 minute, 50_C for 2 minutes and 72°C for 3 minutes, finally 1 cycle at 72_C for 10 minutes.

The PCR amplification with OL1 and OL2 results in a 456 bp fragment (ecycP1, SEQ ID NO: 5), the PCR amplification with OL1

and OL3 results in a 543 bp fragment (ecycP2, SEQ ID NO:_6). The amplicons ecycP1 and ecycP2 are cloned using standard methods into the PCR cloning vector pCR2.1 (Invitrogen), and the clones pTA-ecycP1 and pTA-ecycP2 are obtained. Sequencings of the two clones confirms sequences which are identical in their respective overlap region to SEQ ID NO: 47 and SEQ ID NO: 58, respectively. These clones are therefore used for ligation into the transformation vector pS0301 (WO 02/00900).

The transformation plasmid pEcycP1:GUS is produced by isolating the 454 bp XhoI-NcoI ecycP1 fragment from pTA-ecycP1 and ligating into the XhoI-NcoI-cut vector pS0301. The clone containing the ecycP1 fragment in the correct orientation is called pEcycP1:GUS (Fig.2, construct map).

The transformation plasmid pEcycP2:GUS is produced by isolating the 541 bp XhoI-NcoI ecycP2 fragment from pTA-ecycP2 and ligating into the XhoI-NcoI-cut vector pS0301. The clone containing the ecycP2 fragment in the correct orientation is called pEcycP2:GUS (Fig.3, construct map).

The transformation plasmid pEcycP2:KETO is produced by replacing the region ''GUSI/intron/GUSII/35ST'' bonded by an NcoI and a HindIII restriction cleavage site in pEcycP2:GUS by a ''ketolase/35S terminator'' region. For this purpose, the plasmid pEcycP2:GUS is linearized by standard methods with HindIII, and the resulting 5' overhangs are filled in with Klenow fragment and finally the ''GUSI/intron/GUSII/35ST'' region is deleted by restriction digestion with NcoI.

The ''ketolase/35S terminator'' region is produced by

1. cloning a ketolase cDNA produced using RNA isolated from *Haematococcus pluvialis* (Flotow em. Wille), followed by
2. producing a transcriptional ketolase/terminator fusion by ligating the ketolase sequence into the vector pJIT117, which then serves as template for
3. the PCR amplification of the ketolase/35S terminator region.

The cDNA which codes for the *Haematococcus pluvialis* ketolase is amplified by PCR from *Haematococcus pluvialis* (strain 192.80 of the ''Sammlung von Algenkulturen der Universität Göttingen'') suspension culture.

To prepare total RNA from a suspension culture of *Haematococcus pluvialis* (strain 192.80) which is grown for two weeks in

indirect daylight at room temperature in *Haematococcus* medium (1.2 g/l sodium acetate, 2 g/l yeast extract, 0.2 g/l $MgCl_2 \times 6 H_2O$, 0.02 $CaCl_2 \times 2 H_2O$; pH 6.8; addition of 400 mg/l L-asparagine, 10 mg/l $FeSO_4 \times H_2O$ after autoclaving), the cells are harvested, frozen in liquid nitrogen and powdered in a mortar. 100 mg of the frozen, powdered algal cells are then transferred into a reaction vessel and taken up in 0.8 ml of Trizol® buffer (LifeTechnologies). The suspension is extracted with 0.2 ml of chloroform. After centrifugation at 12 000 g for 15 minutes, the aqueous supernatant is removed and transferred into a new reaction vessel and extracted with one volume of ethanol. The RNA is precipitated with one volume of isopropanol and washed with 75% ethanol, and the pellet is dissolved in DEPC water (water incubated with 1/1000 volume of diethyl pyrocarbonate at room temperature overnight and then autoclaved). The RNA concentration is determined by photometry.

For the cDNA synthesis, 2.5 µg of total RNA are denatured at 60_C for 10 min, cooled on ice for 2 min and transcribed into cDNA using a cDNA kit (Ready-to-go-you-prime-beads®, Pharmacia Biotech) as stated by the manufacturer using an antisense-specific primer (PR1 SEQ ID NO: 64).

The nucleic acid encoding a ketolase from *Haematococcus pluvialis* (strain 192.80) is amplified by the polymer chain reaction (PCR) from *Haematococcus pluvialis* cDNA using a sense-specific primer (PR2; SEQ ID NO: 65) and an antisense-specific primer (PR1; SEQ ID NO: 64). The PCR conditions are as follows:

The PCR for amplifying the cDNA which codes for a ketolase protein consisting of the complete primary sequence takes place in a 50 µl reaction mixture containing:

-	4 µl	a *Haematococcus pluvialis* cDNA (prepared as described above)
-	0.25 µM	dNTPs
-	0.2 µM	primer PR1 (SEQ ID NO: 64)
-	0.2 µM	primer PR2 (SEQ ID NO: 65)
-	5 µl	10X PCR buffer (TAKARA)
-	0.25 µl	R Taq polymerase (TAKARA)
-	25.8 µl	sterile distilled water

The PCR is carried out under the following cycle conditions: 1 cycle at 94_C for 2 minutes; 35 cycles at 94_C for 1 minute,

53_C for 2 minutes and 72_C for 3 minutes. Finally 1 cycle at 72_C for 10 minutes.

The PCR amplification with PR1 and PR2 results in an 1155 bp fragment which codes for a protein consisting of the complete primary sequence. Standard methods are used to clone the ketolase amplicon into the PCR cloning vector pGEM-Teasy (Promega), and the clone pGKETO2 is obtained.

Sequencing of the clone pGKETO2 with the T7 primer and the SP6 primer confirms a sequence which differs only in the three codons 73, 114 and 119, in one base in each case, from the published sequence (Genbank Acc. No.: X86782). These nucleotide exchanges are produced in an independent amplification experiment and thus represent the nucleotide sequence in the Haematococcus pluvialis strain 192.80 used (Fig. 9 and 10, sequence comparisons). This clone is used for cloning into the expression vector pJIT117 (Guerineau et al. (1988) Nucl Acids Res 16: 11380). Further cloning takes place by isolating the 1031 bp SpHI fragment from pGKETO2 and ligating into the SpHI cut vector pJIT117. The clone containing the *Haematococcus pluvialis* ketolase in the correct orientation as N-terminal translational fusion with the rbcs transitpeptide is called pJKETO2.

The 1795 bp ketolase/35S terminator region is produced by PCR using pJKETO2 and the primers OL4 (SEQ ID NO: 62) and OL5 (SEQ ID NO: 63). The conditions of the PCR reactions are as follows:

The PCR for amplifying the OL4-OL5 DNA fragment which contains the coding region of the ketolase followed by the 35S terminator from CaMV takes place in a 50 ∝l reaction mixture containing:

- 1 ∝l pJKETO2 (1 ng of plasmid DNA)
- 0.25 ∝M dNTPs
- 0.2 ∝M primer OL4 (SEQ ID NO: 62)
- 0.2 ∝M primer OL5 (SEQ ID NO: 63)
- 5 ∝l 10X PCR buffer (TAKARA)
- 0.25 ∝l R Taq polymerase (TAKARA)
- 28.8 ∝l sterile distilled water

The PCR reactions are carried out under the following cycle conditions: 1 cycle at 94_C for 2 minutes. 35 cycles at 94_C for 1 minute, 53_C for 2 minutes and 72_C for 3 minutes. Finally 1 cycle at 72_C for 10 minutes.

The PCR amplification with primer OL4 and OL5 results in a 1795 bp fragment which contains the coding region of the ketolase followed by the 35S terminator from CaMV. This 1795 bp amplicon is cloned by using standard methods into the PCR cloning vector pCR2.1 (Invitrogen), and the clone ''pTA-KETO/Term'' is obtained. Sequencing of the clone confirms a sequence which is identical in the respective overlap region to SEQ ID NO: 66 and pJIT117. This clone is therefore used for the ligation into the transformation vector pEcycP2:GUS (see above). The transformation plasmid pEcycP2:KETO is produced by isolating the 1791 bp NcoI-EcoRV ''KETO/Term'' fragment from pTA-KETO/Term and ligating it into the linearized vector pEcycP2:GUS containing an NcoI-5' overhang and a blunt end. The clone containing the ecycP2 fragment in the correct orientation is called pEcycP2:KETO (Fig. 4,construct map).

Example 4: Production and analysis of transgenic tomato plants

The constructs pEcycP1:GUS, pEcycP2:GUS and pEcycP2:KETO were transformed by Agrobacterium tumefaciens-mediated transformation into tomato. Cotyledons and hypocotyls of seedlings seven to ten days old of the Microtom line are used as initial explant for the transformation. The culture medium of Murashige and Skoog (Murashige & Skoog (1962) Physiol Plant 15,473-497) with 2% sucrose, pH 6.1, is used for germination. Germination takes place at 21_C with low light (20 to 100 ∝E). After seven to ten days, the cotyledons are divided transversely, and the hypocotyls are cut into sections about 5 to 10 mm long and placed on the MSBN medium (MS, pH 6.1, 3% sucrose with 1 mg/l benzylaminopurine (BAP), 0.1 mg/l naphthaleneacetate (NAA)) which have been charged the previous day with suspension-cultivated tomato cells. The tomato cells are covered, free of air bubbles, with sterile filter paper. The explants are precultured on the described medium for three to five days. The explants are then infected with the Agrobacterium tumefaciens strain LBA4404, which harbors the binary plasmid with the gene to be transformed, as follows: the strain, which has been cultivated in YEB medium with the antibiotic for the binary plasmid at 28_C overnight, is centrifuged. The bacterial pellet is resuspended in liquid MS medium (3% sucrose, pH 6.1) and adjusted to an optical density of 0.3 (at 600 nm). The precultured explants are transferred into the suspension and incubated at room temperature, shaking gently, for 30 minutes. The explants are then dried with sterile filter paper and returned to their preculture medium for the three-day coculture (21_C).

After the coculture, the explants are transferred to MSZ2 medium (MS pH 6.1 with 3% sucrose, 2 mg/l zeatin, 100 mg/l kanamycin, 160 mg/l timentin) and stored for the selective regeneration at 21_C under weak light conditions (20 to 100 ∝E, light/dark rhythm 16h/8h). The explants are transferred every two to three weeks until shoots form. Small shoots can be detached from the explant and rooted on MS (pH 6.1 with 3% sucrose), 160 mg/l timentin, 30 mg/l kanamycin, 0.1 mg/l IAA. Rooted plants are transferred into the glasshouse.

The transgenicity of rooted tomato plants is confirmed by PCR using genomic DNA. The activity profile of the ε-cyclase promoter fragment can be investigated in the case of the ecycP:GUS construct by a GUS assay by standard methods (Jefferson et al. (1987) EMBO J 6:3901-3907). The activity profile of the ε-cyclase promoter fragment can be investigated in the case of the pEcycP2:KETO construct by Northern blot analysis by standard methods using a ketolase-specific hybridization probe or by ketolase-specific real-time PCR (Sambrook & Russel, 2001, Molecular Cloning: A laboratory manual, 3rd Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York).

Example 5: Production of a transgenic expression vector for producing double-stranded ε-cyclase ribonucleic acid sequences

Expression of inverted repeat transcripts consisting of fragments of the ε-cyclase promoter in *Tagetes erecta* takes place under the control of a modified version (AP3P) of the flower-specific promoter AP3 from *Arabidopsis thaliana* (GenBank Acc. No.: AL132971: Nucleotide region 9298 to 10200; Hill et al. (1998) Development 125:1711-1721). The inverted repeat transcript comprises in each case a fragment in the correct orientation (sense fragment) and a sequence-identical fragment in the contrary orientation (antisense fragment), which are connected together by a functional intron, the PIV2 intron of the potato ST-LH1 gene (Vancanneyt G et al. (1990) Mol Gen Genet 220:245-50).

The cDNA which codes for the AP3 promoter (-902 to +15) from *Arabidopsis thaliana* is produced by PCR using genomic DNA (isolated from *Arabidopsis thaliana* by a standard method) and the primers PR7 (SEQ ID NO: 67) and PR10 (SEQ ID NO: 70). The PCR conditions are as follows:

The PCR for amplifying the DNA which encodes the AP3 promoter fragment (-902 to +15) takes in place in a 50 µl reaction mixture containing:

- 1 µl (equivalent to 20 ng) of genomic DNA from *A.thaliana* (1:100 dil..; prepared as described above)
- 0.25 mM dNTPs
- 0.2 µM primer PR7 (SEQ ID NO: 67)
- 0.2 µM primer PR10 (SEQ ID NO: 70)
- 5 µl 10X PCR buffer (Stratagene)
- 0.25 µl Pfu polymerase (Stratagene)
- 28.8 µl sterile distilled water.

The PCR is carried out under the following cycle conditions: 1 cycle at 94_C for 2 minutes. 35 cycles at 94_C for 1 minute, 50_C for 1 minute and 72_C for 1 minute. Finally 1 cycle at 72_C for 10 minutes.

The 922 bp amplicon is cloned using standard methods into the PCR cloning vector pCR 2.1 (Invitrogen), and the plasmid pTAP3 is obtained. Sequencing of the clone pTAP3 confirms a sequence which differs merely by an insertion (a G in position 9765 of the GenBank Acc. No.: AL132971 sequence) and a base exchange (a G in place of an A in position 9726 of the GenBank Acc. No.: AL132971 sequence) from the published AP3 sequence (GenBank Acc. No.: AL132971, nucleotide region 9298 to 10200) (position 33: T instead of G, position 55: T instead of G). These nucleotide differences can be reproduced in an independent amplification experiment and thus represent the nucleotide sequence in the *Arabidopsis thaliana* plant used.

The modified version AP3P is produced by recombinant PCR using the plasmid pTAP3. The region 10200 to 9771 is amplified using the primers PR7 (SEQ ID NO: 67) and PR9 (SEQ ID NO: 69) (amplicon A7/9), and the region 9526 to 9285 was amplified with the primers PR8 (SEQ ID NO: 68) and PR10 (SEQ ID NO: 70) (amplicon A8/10). The PCR conditions are as follows:

The PCR reaction for amplifying the DNA fragments which code for the regions 10200 to 9771 and 9526 to 9285 of the AP3 promoter takes place in 50 µl reaction mixtures containing:

- 100 ng AP3 amplicon (described above)
- 0.25 mM dNTPs

- 0.2 ∝M primer PR7 (SEQ ID NO: 67) or primer PR8 (SEQ ID NO: 68)
- 0.2 ∝M primer PR9 (SEQ ID NO: 69) or primer PR10 (SEQ ID NO: 70)
- 5 ∝l 10 X PCR buffer (Stratagene)
- 0.25 ∝l Pfu Taq polymerase (Stratagene)
- 28.8 ∝l sterile distilled water

The PCR is carried out under the following cycle conditions: 1 cycle at 94_C for 2 minutes. 35 cycles at 94_C for 1 minute, 50_C for 2 minutes and 72_C for 3 minutes. Finally 1 cycle at 72_C for 10 minutes.

The recombinant PCR comprises annealing of the amplicons A7/9 and A8/10 which overlap over a sequence of 25 nucleotides, completion of a double strand and subsequence amplification. This results in a modified version of the AP3 promoter (AP3P) in which the positions 9670 to 9526 are deleted. Denaturation (5 min at 95_C) and annealing (slow cooling at room temperature to 40_C) of the two amplicons A7/9 and A8/10 takes place in a 17.6 ∝l reaction mixture containing:

- 0.5 ∝g A7/9
- 0.25 ∝g A8/10

Filling-in of the 3' ends (30 min at 30°C), takes place in a 20 ∝l reaction mixture containing:

- 17.6 ∝l A7/9 and A8/10 annealing reaction (prepared as described above)
- 50 ∝M dNTPs
- 2 ∝l 1 X Klenow buffer
- 2 U Klenow enzyme

The nucleic acid coding for the modified promoter version AP3P is amplified by PCR using a sense-specific primer (PR7 SEQ ID NO: 67) and an antisense-specific primer (PR10 SEQ ID NO: 70). The PCR conditions are as follows:

The PCR for amplifying the AP3P fragment takes place in a 50 ∝l reaction mixture containing:

- 1 ∝l annealing reaction (prepared as described above)
- 0.25 mM dNTPs

- 0.2 ∝M primer PR7 (SEQ ID NO: 67)
- 0.2 ∝M primer PR10 (SEQ ID NO: 70)
- 5 ∝l 10 X PCR buffer (Stratagene)
- 0.25 ∝l Pfu Taq polymerase (Stratagene)
- 28.8 ∝l sterile distilled water

The PCR is carried out under the following cycle conditions: 1 cycle at 94_C for 2 minutes. 35 cycles at 94_C for 1 minute, 50_C for 1 minute and 72_C for 1 minute. Finally 1 cycle at 72_C for 10 minutes.

The PCR amplification with the primers PR7 (SEQ ID NO: 67) and PR10 (SEQ ID NO: 70) results in a 777 bp fragment which codes for the modified promoter version AP3P (SEQ ID NO: 77). The amplicon is cloned into the cloning vector pCR2.1 (Invitrogen). Sequencings with the primers T7 and M13 confirm a sequence identical to the sequence of GenBank Acc. No.: AL132971, region 10200 to 9298, the internal region from 9285 to 9526 being deleted. This clone is used for cloning into the expression vector pJIT117 (Guerineau et al. (1988) Nucl Acids Res 16:11380).

The cloning takes place by isolating the 775 bp SacI-HindIII fragment from pTAP3P and ligating into the SacI-HindIII-cut vector pJIT117. The clone which contains the promoter AP3P in place of the original promoter d35S is called pJAP3P.

A DNA fragment which comprises the PIV2 intron of the ST-LS1 gene is produced by PCR using p35SGUS INT plasmid DNA (Vancanneyt G. et al.(1990) Mol Gen Genet 220:245-250) and the primers PR40 (SEQ ID NO: 71) and PR41 (SEQ ID NO: 72). The PCR conditions are as follows:

The PCR for amplifying the sequence of the PIV2 intron of the ST-LS1 gene takes place in a 50 ∝l reaction mixture containing:

- 50 ng p35SGUS INT
- 0.25 mM dNTPs
- 0.2 ∝M primer PR40 (SEQ ID NO: 71)
- 0.2 ∝M primer PR41 (SEQ ID NO: 72)
- 5 ∝l 10X PCR buffer (TAKARA)
- 0.25 ∝l R Taq polymerase (TAKARA)
- 28.8 ∝l sterile distilled water

The PCR is carried out under the following cycle conditions: 1 cycle at 94_C for 2 minutes. 35 cycles at 94_C for 1 minute,

53_C for 1 minute and 72_C for 1 minute. Finally 1 cycle at 72_C for 10 minutes.

The PCR amplification with PR40 and PR41 results in a 212 bp fragment (SEQ ID NO: 78). The amplicon is cloned using standard methods into the PCR cloning vector pBluntII (Invitrogen), and the clone pBluntII-40-41 is obtained. Sequencing of this clone with the primer SP6 confirms a sequence which is identical to the corresponding sequence from the vector p35SGUS INT. This clone is employed for cloning into the vector pJAP3P (see above). The cloning takes place by isolating the 210 bp SalI-BamHI fragment from pBluntII-40-41 and ligating with the SalI-BamHI-cut vector pJAP3P. The clone which contains the PIV2 intron of the ST-LS1 gene in the correct orientation following the 3' end of the rbcs transit peptide is called pJAI1 and is suitable for producing expression cassettes for the flower-specific expression of inverted repeat transcripts.

Example 6: Production of inverted repeat expression cassettes for flower-specific expression of ε-cyclase promoter dsRNAs in *Tagetes erecta*

Expression of inverted repeat transcripts consisting of ε-cyclase promoter fragments in *Tagetes erecta* took place under the control of a modified version (AP3P) of the flower-specific promoter AP3 from *Arabidopsis* (see example 5) or of the flower-specific promoter CHRC (Genbank Acc. No. AF099501). The inverted repeat transcript contains in each case an ε-cyclase promoter fragment in the correct orientation (sense fragment) and a sequence-identical ε-cyclase promoter fragment in the contrary orientation (antisense fragment), which are connected together by a functional intron (see example 5).

The promoter fragments are produced by PCR using plasmid DNA (clone pTA-ecycP, see example 2) and the primers PR124 (SEQ ID NO: 73) and PR126 (SEQ ID NO: 75) or the primers PR125 (SEQ ID NO: 74) and PR127 (SEQ ID NO: 76). The conditions of the PCR reactions are as follows:

The PCR for amplifying the PR124-PR126 DNA fragment which contains the ε-cyclase promoter fragment takes place in a 50 ∝l reaction mixture containing:

- 1 ∝l pTA-ecycP (10 ng/∝l; see example 2)
- 0.25 mM dNTPs

- 0.2 ∝M primer PR124 (SEQ ID NO: 73)
- 0.2 ∝M primer PR126 (SEQ ID NO: 75)
- 5 ∝l 10X PCR buffer (Stratagene)
- 0.25 ∝l Pfu polymerase (Stratagene)
- 28.8 ∝l sterile distilled water

The PCR for amplifying the PR125-PR127 DNA fragment containing the 312 bp ε-cyclase promoter fragment takes place in a 50 ∝l reaction mixture containing:

- 1 ∝l pTA-ecycP (10 ng/∝l; see example 2)
- 0.25 mM dNTPs
- 0.2 ∝M primer PR125 (SEQ ID NO: 74)
- 0.2 ∝M primer PR127 (SEQ ID NO: 76)
- 5 ∝l 10X PCR buffer (Stratagene)
- 0.25 ∝l Pfu polymerase (Stratagene)
- 28.8 ∝l sterile distilled water

The PCR reactions are carried out under the following cycle conditions: 1 cycle at 94_C for 2 minutes. 35 cycles at 94_C for 1 minute, 53_C for 1 minute and 72_C for 1 minute. Finally 1 cycle at 72_C for 10 minutes.

The PCR amplification with primer PR124 and PR126 results in a 358 bp fragment, and PCR amplification with primer PR125 and PR127 resulted in a 361 bp fragment.

The two amplicons, the PR124-PR126 (HindIII-SalI sense) fragment and the PR125-PR127 (EcoRI-BamHI antisense) fragment, are cloned using standard methods into the PCR cloning vector pCR-BluntII (Invitrogen). Sequencings with the primer SP6 confirm in each case a sequence which, apart from the introduced restriction sites, is identical to SEQ ID NO: 58. These clones are therefore used to produce an inverted repeat construct in the cloning vector pJAI1 (see example 5).

The first cloning step takes place by isolating the 356 bp PR124-PR126 HindIII-SalI fragment from the cloning vector pCR-BluntII (Invitrogen) and ligating with the HindIII-SalI-cut vector pJAI1. The clone containing the ε-cyclase promoter fragment in the sense orientation is called cs43. The ligation results in the sense fragment of the ε-cyclase promoter being inserted between the AP3P promoter and the intron. The second cloning step takes place by isolating the 359 bp PR125-PR127 BamHI-EcoRI fragment from the cloning vector pCR-BluntII (Invitrogen) and ligating with BamHI-

EcoRI-cut vector cs43. The clone containing the ε-cyclase promoter fragment in the antisense orientation is called cs44. The ligation results in a transcriptional fusion between the intron and the antisense fragment of the ε-cyclase promoter.

An inverted repeat expression cassette under the control of the CHRC promoter is produced by amplifying a CHRC promoter fragment using genomic petunia DNA (prepared by standard methods) and the primers PRCHRC5' (SEQ ID NO 82) and PRCHRC3' (SEQ ID NO: 83). The amplicon is cloned into the cloning vector pCR2.1 (Invitrogen). Sequencings of the resulting clone pCR2.1-CHRC with the primers M13 and T7 confirm a sequence identical to the GenBank Acc. No.: AF099501 sequence. This clone is therefore used for cloning into the expression vector cs44. The cloning takes place by isolating the 1535 bp SacI-HindIII fragment from pCR2.1-CHRC and ligating into the SacI-HindIII-cut vector cs44. The clone which contains the CHRC promoter in place of the original AP3P promoter is called cs45.

The transformation plasmids for the agrobacterium-mediated transformation of the AP3P-controlled inverted repeat transcript in Tagetes erecta are produced using the binary vector pSUN5 (WO 02/00900).

The transformation plasmid pS5AI7 is produced by ligating the 1683 bp SacI-XhoI fragment from cs44 with the SacI-XhoI-cut vector pSUN5 (Fig. 5, construct map).

The transformation plasmid pS5CI7 is produced by ligating the 2448 bp SacI-XhoI fragment from cs45 with the SacI-XhoI-cut vector pSUN5 (Fig.6, construct map).

Example 7: Production and analysis of transgenic Tagetes plants

The transformation plasmids pS5AI7 and pS5CI7 are transformed by Agrobacterium tumefaciens-mediated transformation into Tagetes.

Tagetes seeds are sterilized and placed on germination medium (MS medium; Murashige & Skoog (1962) Physiol Plant 15:473-497; pH 5.8, 2% sucrose). Germination takes place in a temperture/light/time interval of 18 to 28_C/20 to 200 ∝E/3 to 16 weeks, but preferably at 21_C, 20 to 70 ∝E, for 4 to 8 weeks.

All leaves of the plants which have developed in vitro by then are harvested and cut transverse to the middle. The leaf explants resulting therefrom, with a size of 10 to 60 mm^2, are stored

during the preparation in liquid MS medium at room temperature for not more than 2 h.

Any Agrobacterium tumefaciens strain, but preferably a supervirulent strain such as, for example, EHA105 with an appropriate binary plasmid, which may harbor a selection marker gene (preferably *bar* or *pat*) and one or more trait or reporter genes is cultivated overnight and used for cocultivation with the leaf material. The bacterial strain can be cultured as follows: a single colony of the appropriate strain is inoculated in YEB (0.1% yeast extract, 0.5% beef extract, 0.5% peptone, 0.5% sucrose, 0.5% magnesium sulfate x 7 H_2O) with 25 mg/l kanamycin and cultured at 28_C for 16 to 20 h. The bacterial suspension is then harvested by centrifugation at 6000 g for 10 min, and resuspended in liquid MS medium so as to result in an OD_{600} of about 0.1 to 0.8.

Immediately before the cocultivation, the MS medium in which the leaves have been stored is replaced by the bacterial suspension. Incubation of the leaves in the agrobacteria suspension took place at room temperature with gentle shaking for 30 min. The infected explants are then put on an MS medium solified with agar (e.g. 0.8% plant agar (Duchefa, NL)), with growth regulators such as, for example, 3 mg/l benzylaminopurine (BAP) and 1 mg/l indolylacetic acid (IAA). The orientation of the leaves on the medium is immaterial. Cultivation of the explants takes place for 1 to 8 days, but preferably for 6 days, during which the following conditions can be used: light intensity: 30 to 80 $\mu mol/m^2 \times sec$, temperature: 22 to 24°C, 16/18 hours light/dark alternation. The cocultivated explants are then transferred to fresh MS medium, preferably with the same growth regulators, this second medium additionally containing an antibiotic to suppress bacterial growth. Timentin in a concentration of 200 to 500 mg/l is very suitable for this purpose. The second selective component employed is one for selecting for successful transformation. Phosphinothricin in a concentration of 1 to 5 mg/l selects very efficiently, but other selective components according to the method to be used are also conceivable. After one to three weeks in each case, the explants are transferred to fresh medium until plumules and small shoots develop, which are then transferred to the same basal medium including timentin and PPT or alternative components with growth regulators, namely, for example, 0.5 mg/l indolylbutyric acid (IBA) and 0.5 mg/l gibberillic acid GA_3, for rooting. Rooted shoots can be transferred into the glasshouse.

In addition to the method described, the following advantageous modifications are possible:

- before the explants are infected with the bacteria, they can be preincubated on the medium described above for the cocultivation for 1 to 12 days, preferably 3 to 4. This is followed by infection, cocultivation and selection regeneration as described above.

- the pH for the regeneration (normally 5.8) can be lowered to pH 5.2. This improves control of the growth of agrobacteria.

- addition of $AgNO_3$ (3 to 10 mg/l) to the regeneration medium improves the condition of the culture, including the regeneration itself.

- components which reduce phenol formation and are known to the skilled worker, such as, for example citric acid, ascorbic acid, PVP and many others, have beneficial effects on the culture.

- liquid culture medium can also be used for the whole method. The culture can also be incubated on commercially available supports which are positioned on the liquid medium.

The transgenicity of rooted shoots can be investigated on isolated genomic DNA by the polymerase chain reaction (PCR). The reduction in the amounts of ε-cyclase transcript (compared with the wild type used for the transformation) as a result of transformation with the transformation plasmid pS5AI7 or pS5CI7 can be investigated by Northern gel blot analysis by standard methods (Sambrook & Russel, 2001, Molecular Cloning: A laboratory manual, 3rd Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York) using an ε-cyclase-specific hybridization probe, for example produced as described in example 1. In addition, the reduction in the amounts of ε-cyclase transcript (compared with the wild type used for the transformation) can be investigated by ε-cyclase-specific real time PCR.

SEQUENCE LISTING

<110> SunGene GmbH & Co.KGaA

<120> Transgenic expression cassettes for expressing nucleic acids in the flower of plants

<130> AE20020666

<140>
<141>

<160> 83

<170> PatentIn Ver. 2.1

<210> 1
<211> 312
<212> DNA
<213> Tagetes erecta

<220>
<221> promoter
<222> (1)..(312)

<400> 1
cgatagtaaa atcgttagtt atgattaata cttgggaggt gggggattat aggctttgtt 60
gtgagaatgt tgagaaagag gtttgacaaa tcggtgtttg aatgaggtta aatggagttt 120
aattaaaata aagagaagag aaagattaag agggtgatgg ggatattaaa gacggscaat 180
atagtgatgc cacgtagaaa aaggtaagtg aaaacataca acgtggcttt aaaagatggc 240
ttggctgcta atcaactcaa ctcaactcat atcctatcca ttcaaattca attcaattct 300
attgaatgca aa 312

<210> 2
<211> 447
<212> DNA
<213> Tagetes erecta

<220>
<221> promoter
<222> (1)..(312)

<220>
<221> 5'UTR
<222> (313)..(444)

<220>
<221> misc_feature
<222> (445)..(447)
<223> ATG-start codon

<400> 2
cgatagtaaa atcgttagtt atgattaata cttgggaggt gggggattat aggctttgtt 60
gtgagaatgt tgagaaagag gtttgacaaa tcggtgtttg aatgaggtta aatggagttt 120
aattaaaata aagagaagag aaagattaag agggtgatgg ggatattaaa gacggscaat 180
atagtgatgc cacgtagaaa aaggtaagtg aaaacataca acgtggcttt aaaagatggc 240
ttggctgcta atcaactcaa ctcaactcat atcctatcca ttcaaattca attcaattct 300
attgaatgca aagcaaagca aaggttgttt gttgttgttg ttgagagaca ctccaatcca 360
aacagataca aggcgtgact ggatatttct ctctcgttcc taacaacagc aacgaagaag 420
aaaaagaatc attactaaca atcaatg 447

<210> 3
<211> 537
<212> DNA
<213> Tagetes erecta

<220>
<221> promoter
<222> (1)..(312)

<220>
<221> 5'UTR
<222> (313)..(444)

<220>
<221> sig_peptide
<222> (445)..(537)

<220>
<221> CDS
<222> (445)..(537)
<223> coding for transit peptide of epsilon cyclase

<400> 3
cgatagtaaa atcgttagtt atgattaata cttgggaggt gggggattat aggctttgtt 60
gtgagaatgt tgagaaagag gtttgacaaa tcggtgtttg aatgaggtta aatggagttt 120
aattaaaata aagagaagag aaagattaag agggtgatgg ggatattaaa gacggscaat 180
atagtgatgc cacgtagaaa aaggtaagtg aaaacataca acgtggcttt aaaagatggc 240
ttggctgcta atcaactcaa ctcaactcat atcctatcca ttcaaattca attcaattct 300
attgaatgca aagcaaagca aaggttgttt gttgttgttg ttgagagaca ctccaatcca 360
aacagataca aggcgtgact ggatatttct ctctcgttcc taacaacagc aacgaagaag 420
aaaaagaatc attactaaca atca atg agt atg aga gct gga cac atg acg 471
Met Ser Met Arg Ala Gly His Met Thr
1 5

gca aca atg gcg gct ttt aca tgc cct agg ttt atg act agc atc aga 519
Ala Thr Met Ala Ala Phe Thr Cys Pro Arg Phe Met Thr Ser Ile Arg
10 15 20 25

tac acg aag caa att aag 537
Tyr Thr Lys Gln Ile Lys
30

<210> 4
<211> 31
<212> PRT
<213> Tagetes erecta

<400> 4
Met Ser Met Arg Ala Gly His Met Thr Ala Thr Met Ala Ala Phe Thr
1 5 10 15
Cys Pro Arg Phe Met Thr Ser Ile Arg Tyr Thr Lys Gln Ile Lys
20 25 30

<210> 5
<211> 456
<212> DNA
<213> Tagetes erecta

<220>
<221> misc_feature
<222> (1)..(6)
<223> restriction site

<220>
<221> misc_feature
<222> (451)..(456)
<223> restriction site

<220>
<221> misc_feature
<222> (7)..(450)
<223> coding for promoter and 5'-UTR region

<400> 5
ctcgagagta aaatcgttag ttatgattaa tacttgggag gtgggggatt ataggctttg 60
ttgtgagaat gttgagaaag aggtttgaca aatcggtgtt tgaatgaggt taaatggagt 120
ttaattaaaa taaagagaag agaaagatta agagggtgat ggggatatta aagacggsca 180
atatagtgat gccacgtaga aaaaggtaag tgaaaacata caacgtggct ttaaaagatg 240
gcttggctgc taatcaactc aactcaactc atatcctatc cattcaaatt caattcaatt 300
ctattgaatg caaagcaaag caaaggttgt ttgttgttgt tgttgagaga cactccaatc 360
caaacagata caaggcgtga ctggatattt ctctctcgtt cctaacaaca gcaacgaaga 420
agaaaaagaa tcattactaa caatcaatgg ccatgg 456

<210> 6
<211> 543
<212> DNA
<213> Tagetes erecta

<220>
<221> misc_feature
<222> (1)..(6)
<223> restriction site

<220>
<221> misc_feature
<222> (538)..(543)
<223> restriction site

<220>
<221> misc_feature
<222> (7)..(542)
<223> coding for promoter - 5'-UTR - signal peptide region

<400> 6
ctcgagagta aaatcgttag ttatgattaa tacttgggag gtgggggatt ataggctttg 60
ttgtgagaat gttgagaaag aggtttgaca aatcggtgtt tgaatgaggt taaatggagt 120
ttaattaaaa taaagagaag agaaagatta agagggtgat ggggatatta aagacggsca 180
atatagtgat gccacgtaga aaaaggtaag tgaaaacata caacgtggct ttaaaagatg 240
gcttggctgc taatcaactc aactcaactc atatcctatc cattcaaatt caattcaatt 300
ctattgaatg caaagcaaag caaaggttgt ttgttgttgt tgttgagaga cactccaatc 360
caaacagata caaggcgtga ctggatattt ctctctcgtt cctaacaaca gcaacgaaga 420
agaaaaagaa tcattactaa caatcaatga gtatgagagc tggacacatg acggcaacaa 480
tggcggcttt tacatgccct aggtttatga ctagcatcag atacacgaag caaattacca 540
tgg 543

<210> 7
<211> 1008
<212> DNA
<213> Arabidopsis thaliana

<400> 7
gaagatgtcg tggacagtta cctagtagaa agccctgaaa ctcacgatgt cactgacttc 60
tgcagcttct acactctccc ttcaaccatc ctcggtaacc cgaactacac tacattgaaa 120
gctgcgtatt cttactacaa tgtggccaca cagacctcgt ttcttcagct gatgaatgat 180

gcgctaattg tctcaaagca aaagggtttc gatgtgttca acgcgttgga tgtgatgcac 240
aatgagagtt tcttgaaaga actgaagttt gggccaggag atggacaact tcattactat 300
ctctacaatt accgtttgaa aagtgccttg aagccagcgg aactcgggct tgttctctta 360
taagctcaac aacttgattt gatggtatca acaaacttga aatttgtctc tctttttttt 420
tcttcagtct gaactacttc tcccatggtt tactgaaact gagtttaatt attttggcaa 480
tgcaattttg gattttgctt ttagtttcac ttgtgttctc tgagagagtt gaccctgaga 540
agtaaagtca acactagacg tgtgcgagtt tcaatggact gcatttttag gtccaagaag 600
tagctacggc ccatctatcg tagaagttgt tttttgtact aaagcctgtt atagccgaac 660
taggtcgcta cgtttcctaa cttgtcgttt attaggctac ttttaaatcg cacgcaccaa 720
tccagttctt acacgcgttt tcacaaaaag gaccttacat gggttgttcc cccaataaga 780
aaattacacg ttgcagaatc gaagcagtgg tctatccttg ttatcgtcct ctcgtcttta 840
tggctcttcc aagtcatttt gagaaatttt cgattttgag attttttttcc aaatattaca 900
aaaggaaata attagattcc tcttttctgct tgctatacct tgatagaaca atataacaat 960
ggtgtaagtc ttctcgctgt attcgaaatt atttggagga ggaaaatg 1008

<210> 8
<211> 1169
<212> DNA
<213> Oryza sativa

<400> 8
agaaggagag agagccggcc ggcgcgacgg cggtggcagg agaggtggag agagcccgcc 60
gacgcgacgg cggcagtgga agaaaaggag aggaggcggc cggcgctgtc gaacgcgcac 120
gggaagcagc ggacgatgcc gagcaggaca atgtcctcct cgtcaccacc ttcctcatcg 180
tcgtcgcggt cggtgggtgc cgggcggagc ggtttttgag ggaggggtcg gcggccggcg 240
agcccatccg agctcgcctg gcttgcgcga gcccgccctc gctgcagctt ggtgtgctgg 300
ggacgacgac gaggaaaggg ggagagagga atcaactacc gacgacagcg gaggcgggcg 360
agcgcggtgc cgggctgctc acgctcgccg accaggccct cgctccgtca cactcgcctg 420
cctccgttcc acccgcagcc cgcgccgctc cgttcgcagc gctcgcccgt cgccgctccg 480
ctgctgtgca gtcccgccgc cgcgctcgcc ctgactgaag aagaaaagag agaagagaga 540
aaagagaagg gaaggagaag aaaatagaag aaaaaaatat gtgcagctga tgtatgagcc 600
ccacatactc ttttttaatc ttttttgctg actacgatgc cacgtcagcg aaaccaccta 660
tatatactac cataggatct tgagtgcacg gtttatatga gtttaggagt atacattttt 720
agttttatgg ttaagggatc ataaaaaatt ctcgctatta agttgagtga cgcgcagtga 780
acttattact caaacttaac agcgtttgat ccattcacat ccggcccata gcagcccata 840
tccagcaacg caccacggtg actgacacgc ggggcccaac gaggtggcgc cgtggctggt 900
ggggcgggtg agtcacgccc tcggcgctgc ccccaaaagc gaaccatatc ctttgcttcg 960
tggggccctg cccaatcgcc gcccgccacg tgccccttcc cggataccgc ctccctccaa 1020
acccgctccc tctccctctc cctcctccta caatggccgc agcagcagca gccagcagca 1080
gacgcagaga ccagtagctt tcgcagaggc ggcagccacc accgcctcct cctcctcctc 1140
ctcatccgac ggtgtgcaac caagcgatg 1169

<210> 9
<211> 1830
<212> DNA
<213> Tagetes erecta

<220>
<221> CDS
<222> (141)..(1688)
<223> coding for epsilon-cyclase

<400> 9
ggcacgaggc aaagcaaagg ttgtttgttg ttgttgttga gagacactcc aatccaaaca 60
gatacaaggc gtgactggat atttctctct cgttcctaac aacagcaacg aagaagaaaa 120
agaatcatta ctaacaatca atg agt atg aga gct gga cac atg acg gca aca 173
Met Ser Met Arg Ala Gly His Met Thr Ala Thr
1 5 10

```
atg gcg gct ttt aca tgc cct agg ttt atg act agc atc aga tac acg   221
Met Ala Ala Phe Thr Cys Pro Arg Phe Met Thr Ser Ile Arg Tyr Thr
                15                  20                  25

aag caa att aag tgc aac gct gct aaa agc cag cta gtc gtt aaa caa   269
Lys Gln Ile Lys Cys Asn Ala Ala Lys Ser Gln Leu Val Val Lys Gln
            30                  35                  40

gag att gag gag gaa gaa gat tat gtg aaa gcc ggt gga tcg gag ctg   317
Glu Ile Glu Glu Glu Glu Asp Tyr Val Lys Ala Gly Gly Ser Glu Leu
        45                  50                  55

ctt ttt gtt caa atg caa cag aat aag tcc atg gat gca cag tct agc   365
Leu Phe Val Gln Met Gln Gln Asn Lys Ser Met Asp Ala Gln Ser Ser
    60                  65                  70                  75

cta tcc caa aag ctc cca agg gta cca ata gga gga gga gga gac agt   413
Leu Ser Gln Lys Leu Pro Arg Val Pro Ile Gly Gly Gly Gly Asp Ser
                80                  85                  90

aac tgt ata ctg gat ttg gtt gta att ggt tgt ggt cct gct ggc ctt   461
Asn Cys Ile Leu Asp Leu Val Val Ile Gly Cys Gly Pro Ala Gly Leu
            95                  100                 105

gct ctt gct gga gaa tca gcc aag cta ggc ttg aat gtc gca ctt atc   509
Ala Leu Ala Gly Glu Ser Ala Lys Leu Gly Leu Asn Val Ala Leu Ile
        110                 115                 120

ggc cct gat ctt cct ttt aca aat aac tat ggt gtt tgg gag gat gaa   557
Gly Pro Asp Leu Pro Phe Thr Asn Asn Tyr Gly Val Trp Glu Asp Glu
    125                 130                 135

ttt ata ggt ctt gga ctt gag ggc tgt att gaa cat gtt tgg cga gat   605
Phe Ile Gly Leu Gly Leu Glu Gly Cys Ile Glu His Val Trp Arg Asp
140                 145                 150                 155

act gta gta tat ctt gat gac aac gat ccc att ctc ata ggt cgt gcc   653
Thr Val Val Tyr Leu Asp Asp Asn Asp Pro Ile Leu Ile Gly Arg Ala
                160                 165                 170

tat gga cga gtt agt cgt gat tta ctt cac gag gag ttg ttg act agg   701
Tyr Gly Arg Val Ser Arg Asp Leu Leu His Glu Glu Leu Leu Thr Arg
            175                 180                 185

tgc atg gag tca ggc gtt tca tat ctg agc tcc aaa gtg gaa cgg att   749
Cys Met Glu Ser Gly Val Ser Tyr Leu Ser Ser Lys Val Glu Arg Ile
        190                 195                 200

act gaa gct cca aat ggc cta agt ctc ata gag tgt gaa ggc aat atc   797
Thr Glu Ala Pro Asn Gly Leu Ser Leu Ile Glu Cys Glu Gly Asn Ile
    205                 210                 215

aca att cca tgc agg ctt gct act gtc gct tct gga gca gct tct gga   845
Thr Ile Pro Cys Arg Leu Ala Thr Val Ala Ser Gly Ala Ala Ser Gly
220                 225                 230                 235

aaa ctt ttg cag tat gaa ctt ggc ggt ccc cgt gtt tgc gtt caa aca   893
Lys Leu Leu Gln Tyr Glu Leu Gly Gly Pro Arg Val Cys Val Gln Thr
                240                 245                 250

gct tat ggt ata gag gtt gag gtt gaa agc ata ccc tat gat cca agc   941
Ala Tyr Gly Ile Glu Val Glu Val Glu Ser Ile Pro Tyr Asp Pro Ser
            255                 260                 265
```

```
cta atg gtt ttc atg gat tat aga gac tac acc aaa cat aaa tct caa    989
Leu Met Val Phe Met Asp Tyr Arg Asp Tyr Thr Lys His Lys Ser Gln
        270                 275                 280

tca cta gaa gca caa tat cca aca ttt ttg tat gtc atg cca atg tct   1037
Ser Leu Glu Ala Gln Tyr Pro Thr Phe Leu Tyr Val Met Pro Met Ser
    285                 290                 295

cca act aaa gta ttc ttt gag gaa act tgt ttg gct tca aaa gag gcc   1085
Pro Thr Lys Val Phe Phe Glu Glu Thr Cys Leu Ala Ser Lys Glu Ala
300                 305                 310                 315

atg cct ttt gag tta ttg aag aca aaa ctc atg tca aga tta aag act   1133
Met Pro Phe Glu Leu Leu Lys Thr Lys Leu Met Ser Arg Leu Lys Thr
                320                 325                 330

atg ggg atc cga ata acc aaa act tat gaa gag gaa tgg tca tat att   1181
Met Gly Ile Arg Ile Thr Lys Thr Tyr Glu Glu Glu Trp Ser Tyr Ile
            335                 340                 345

cca gta ggt gga tcc tta cca aat acc gag caa aag aac ctt gca ttt   1229
Pro Val Gly Gly Ser Leu Pro Asn Thr Glu Gln Lys Asn Leu Ala Phe
        350                 355                 360

ggt gct gct gct agc atg gtg cat cca gcc aca gga tat tcg gtt gta   1277
Gly Ala Ala Ala Ser Met Val His Pro Ala Thr Gly Tyr Ser Val Val
    365                 370                 375

aga tca ctg tca gaa gct cct aat tat gca gca gta att gca aag att   1325
Arg Ser Leu Ser Glu Ala Pro Asn Tyr Ala Ala Val Ile Ala Lys Ile
380                 385                 390                 395

tta ggg aaa gga aat tca aaa cag atg ctt gat cat gga aga tac aca   1373
Leu Gly Lys Gly Asn Ser Lys Gln Met Leu Asp His Gly Arg Tyr Thr
                400                 405                 410

acc aac atc tca aag caa gct tgg gaa aca ctt tgg ccc ctt gaa agg   1421
Thr Asn Ile Ser Lys Gln Ala Trp Glu Thr Leu Trp Pro Leu Glu Arg
            415                 420                 425

aaa aga cag aga gca ttc ttt ctc ttt gga tta gca ctg att gtc cag   1469
Lys Arg Gln Arg Ala Phe Phe Leu Phe Gly Leu Ala Leu Ile Val Gln
        430                 435                 440

atg gat att gag ggg acc cgc aca ttc ttc cgg act ttc ttc cgc ttg   1517
Met Asp Ile Glu Gly Thr Arg Thr Phe Phe Arg Thr Phe Phe Arg Leu
    445                 450                 455

ccc aca tgg atg tgg tgg ggg ttt ctt gga tct tcg tta tca tca act   1565
Pro Thr Trp Met Trp Trp Gly Phe Leu Gly Ser Ser Leu Ser Ser Thr
460                 465                 470                 475

gac ttg ata ata ttt gcg ttt tac atg ttt atc ata gca ccg cat agc   1613
Asp Leu Ile Ile Phe Ala Phe Tyr Met Phe Ile Ile Ala Pro His Ser
                480                 485                 490

ctg aga atg ggt ctg gtt aga cat ttg ctt tct gac ccg aca gga gga   1661
Leu Arg Met Gly Leu Val Arg His Leu Leu Ser Asp Pro Thr Gly Gly
            495                 500                 505

aca atg tta aaa gcg tat ctc acg ata taaataactc tagtcgcgat         1708
Thr Met Leu Lys Ala Tyr Leu Thr Ile
        510                 515

cagtttagat tataggcaca tcttgcatat atatatgtat aaaccttatg tgtgctgtat 1768

ccttacatca acacagtcat taattgtatt tcttggggta atgctgatga agtattttct 1828
```

gg 1830

<210> 10
<211> 516
<212> PRT
<213> Tagetes erecta

<400> 10

Met Ser Met Arg Ala Gly His Met Thr Ala Thr Met Ala Ala Phe Thr
1 5 10 15

Cys Pro Arg Phe Met Thr Ser Ile Arg Tyr Thr Lys Gln Ile Lys Cys
20 25 30

Asn Ala Ala Lys Ser Gln Leu Val Val Lys Gln Glu Ile Glu Glu Glu
35 40 45

Glu Asp Tyr Val Lys Ala Gly Gly Ser Glu Leu Leu Phe Val Gln Met
50 55 60

Gln Gln Asn Lys Ser Met Asp Ala Gln Ser Ser Leu Ser Gln Lys Leu
65 70 75 80

Pro Arg Val Pro Ile Gly Gly Gly Gly Asp Ser Asn Cys Ile Leu Asp
85 90 95

Leu Val Val Ile Gly Cys Gly Pro Ala Gly Leu Ala Leu Ala Gly Glu
100 105 110

Ser Ala Lys Leu Gly Leu Asn Val Ala Leu Ile Gly Pro Asp Leu Pro
115 120 125

Phe Thr Asn Asn Tyr Gly Val Trp Glu Asp Glu Phe Ile Gly Leu Gly
130 135 140

Leu Glu Gly Cys Ile Glu His Val Trp Arg Asp Thr Val Val Tyr Leu
145 150 155 160

Asp Asp Asn Asp Pro Ile Leu Ile Gly Arg Ala Tyr Gly Arg Val Ser
165 170 175

Arg Asp Leu Leu His Glu Glu Leu Leu Thr Arg Cys Met Glu Ser Gly
180 185 190

Val Ser Tyr Leu Ser Ser Lys Val Glu Arg Ile Thr Glu Ala Pro Asn
195 200 205

Gly Leu Ser Leu Ile Glu Cys Glu Gly Asn Ile Thr Ile Pro Cys Arg
210 215 220

Leu Ala Thr Val Ala Ser Gly Ala Ala Ser Gly Lys Leu Leu Gln Tyr
225 230 235 240

Glu Leu Gly Gly Pro Arg Val Cys Val Gln Thr Ala Tyr Gly Ile Glu
245 250 255

Val Glu Val Glu Ser Ile Pro Tyr Asp Pro Ser Leu Met Val Phe Met
260 265 270

Asp Tyr Arg Asp Tyr Thr Lys His Lys Ser Gln Ser Leu Glu Ala Gln
275 280 285

Tyr Pro Thr Phe Leu Tyr Val Met Pro Met Ser Pro Thr Lys Val Phe
290 295 300

Phe Glu Glu Thr Cys Leu Ala Ser Lys Glu Ala Met Pro Phe Glu Leu
305 310 315 320

Leu Lys Thr Lys Leu Met Ser Arg Leu Lys Thr Met Gly Ile Arg Ile
325 330 335

```
Thr Lys Thr Tyr Glu Glu Glu Trp Ser Tyr Ile Pro Val Gly Gly Ser
            340                 345                 350
Leu Pro Asn Thr Glu Gln Lys Asn Leu Ala Phe Gly Ala Ala Ala Ser
        355                 360                 365
Met Val His Pro Ala Thr Gly Tyr Ser Val Val Arg Ser Leu Ser Glu
    370                 375                 380
Ala Pro Asn Tyr Ala Ala Val Ile Ala Lys Ile Leu Gly Lys Gly Asn
385                 390                 395                 400
Ser Lys Gln Met Leu Asp His Gly Arg Tyr Thr Thr Asn Ile Ser Lys
                405                 410                 415
Gln Ala Trp Glu Thr Leu Trp Pro Leu Glu Arg Lys Arg Gln Arg Ala
            420                 425                 430
Phe Phe Leu Phe Gly Leu Ala Leu Ile Val Gln Met Asp Ile Glu Gly
        435                 440                 445
Thr Arg Thr Phe Phe Arg Thr Phe Phe Arg Leu Pro Thr Trp Met Trp
    450                 455                 460
Trp Gly Phe Leu Gly Ser Ser Leu Ser Ser Thr Asp Leu Ile Ile Phe
465                 470                 475                 480
Ala Phe Tyr Met Phe Ile Ile Ala Pro His Ser Leu Arg Met Gly Leu
                485                 490                 495
Val Arg His Leu Leu Ser Asp Pro Thr Gly Gly Thr Met Leu Lys Ala
            500                 505                 510
Tyr Leu Thr Ile
        515
```

<210> 11
<211> 1916
<212> DNA
<213> Tagetes erecta

<220>
<221> CDS
<222> (173)..(1747)
<223> coding for epsilon-cyclase

<400> 11

```
attcaaattc aattcaattc tattgaatgc aaagcaaagc aaagcaaagg ttgtttgttg 60
ttgttgtttg agagacactc caatccaaac agatacaagg cgtgactgga tatttctctc 120
tcgttcctta acaacagcaa cgaagaagaa aaagaatcat tactcacaat ca atg agt 178
                                                        Met Ser
                                                          1
atg aga gct gga cac atg acg gca aca atg gcg gct ttt aca tgc cct 226
Met Arg Ala Gly His Met Thr Ala Thr Met Ala Ala Phe Thr Cys Pro
          5                  10                  15
agg ttt atg act agc atc aga tac acg aag caa att aag tgc aac gct 274
Arg Phe Met Thr Ser Ile Arg Tyr Thr Lys Gln Ile Lys Cys Asn Ala
     20                  25                  30
gct aaa agc cag cta gtc gtt aaa caa gag att gag gag gaa gaa gat 322
Ala Lys Ser Gln Leu Val Val Lys Gln Glu Ile Glu Glu Glu Glu Asp
 35                  40                  45                  50
```

```
tat gtg aaa gcc ggt gga tcg gag ctg ctt ttt gtt caa atg caa cag   370
Tyr Val Lys Ala Gly Gly Ser Glu Leu Leu Phe Val Gln Met Gln Gln
                 55                  60                  65

aat aag tcc atg gat gca cag tct agc cta tcc caa aag ctc cca agg   418
Asn Lys Ser Met Asp Ala Gln Ser Ser Leu Ser Gln Lys Leu Pro Arg
             70                  75                  80

gta cca ata gga gga gga gga gac agt aac tgt ata ctg gat ttg gtt   466
Val Pro Ile Gly Gly Gly Gly Asp Ser Asn Cys Ile Leu Asp Leu Val
         85                  90                  95

gta att ggt tgt ggt cct gct ggc ctt gct ctt gct gga gaa tca gcc   514
Val Ile Gly Cys Gly Pro Ala Gly Leu Ala Leu Ala Gly Glu Ser Ala
    100                 105                 110

aag cta ggc ttg aat gtc gca ctt atc ggc cct gat ctt cct ttt aca   562
Lys Leu Gly Leu Asn Val Ala Leu Ile Gly Pro Asp Leu Pro Phe Thr
115                 120                 125                 130

aat aac tat ggt gtt tgg gag gat gaa ttt ata ggt ctt gga ctt gag   610
Asn Asn Tyr Gly Val Trp Glu Asp Glu Phe Ile Gly Leu Gly Leu Glu
                135                 140                 145

ggc tgt att gaa cat gtt tgg cga gat act gta gta tat ctt gat gac   658
Gly Cys Ile Glu His Val Trp Arg Asp Thr Val Val Tyr Leu Asp Asp
            150                 155                 160

aac gat ccc att ctc ata ggt cgt gcc tat gga cga gtt agt cgt gat   706
Asn Asp Pro Ile Leu Ile Gly Arg Ala Tyr Gly Arg Val Ser Arg Asp
        165                 170                 175

tta ctt cac gag gag ttg ttg act agg tgc atg gag tca ggc gtt tca   754
Leu Leu His Glu Glu Leu Leu Thr Arg Cys Met Glu Ser Gly Val Ser
    180                 185                 190

tat ctg agc tcc aaa gtg gaa cgg att act gaa gct cca aat ggc cta   802
Tyr Leu Ser Ser Lys Val Glu Arg Ile Thr Glu Ala Pro Asn Gly Leu
195                 200                 205                 210

agt ctc ata gag tgt gaa ggc aat atc aca att cca tgc agg ctt gct   850
Ser Leu Ile Glu Cys Glu Gly Asn Ile Thr Ile Pro Cys Arg Leu Ala
                215                 220                 225

act gtc gct tct gga gca gct tct ggg aaa ctt ttg cag tat gaa ctt   898
Thr Val Ala Ser Gly Ala Ala Ser Gly Lys Leu Leu Gln Tyr Glu Leu
            230                 235                 240

ggc ggt ccc cgt gtt tgc gtt caa aca gct tat ggt tac gag gtt gag   946
Gly Gly Pro Arg Val Cys Val Gln Thr Ala Tyr Gly Tyr Glu Val Glu
        245                 250                 255

gtt gaa agc ata ccc tat gat cca agc cta atg gtt ttc atg gat tat   994
Val Glu Ser Ile Pro Tyr Asp Pro Ser Leu Met Val Phe Met Asp Tyr
    260                 265                 270

aga gac tac acc aaa cat aaa tct caa tca cta gaa gca caa tat cca  1042
Arg Asp Tyr Thr Lys His Lys Ser Gln Ser Leu Glu Ala Gln Tyr Pro
275                 280                 285                 290

aca ttt ttg tat gtc atg cca atg tct cca act aaa gta ttc ttt gag  1090
Thr Phe Leu Tyr Val Met Pro Met Ser Pro Thr Lys Val Phe Phe Glu
                295                 300                 305
```

```
gaa act tgt ttg gct tca aaa gag gcc atg cct ttt gag tta ttg aag   1138
Glu Thr Cys Leu Ala Ser Lys Glu Ala Met Pro Phe Glu Leu Leu Lys
            310                 315                 320

aca aaa ctc atg tca aga tta aag act atg ggg atc cga ata acc aaa   1186
Thr Lys Leu Met Ser Arg Leu Lys Thr Met Gly Ile Arg Ile Thr Lys
        325                 330                 335

act tat gaa gag tat ctt gtt gct tgt caa tat ttg gaa gaa tgg tca   1234
Thr Tyr Glu Glu Tyr Leu Val Ala Cys Gln Tyr Leu Glu Glu Trp Ser
    340                 345                 350

tat att cca gta ggt gga tcc ctt cca aat acc gag caa aag aac ctt   1282
Tyr Ile Pro Val Gly Gly Ser Leu Pro Asn Thr Glu Gln Lys Asn Leu
355                 360                 365                 370

gca ttt ggt gct gct gct agc atg gtg cat cca gcc aca gga tat tcg   1330
Ala Phe Gly Ala Ala Ala Ser Met Val His Pro Ala Thr Gly Tyr Ser
                375                 380                 385

gtt gta aga tca ctg tca gaa gct cct aat tat gca gca gta att gca   1378
Val Val Arg Ser Leu Ser Glu Ala Pro Asn Tyr Ala Ala Val Ile Ala
            390                 395                 400

aag att tta ggg aaa gga aat tca aaa cag atg ctt gat ctt gga aga   1426
Lys Ile Leu Gly Lys Gly Asn Ser Lys Gln Met Leu Asp Leu Gly Arg
        405                 410                 415

tac aca acc aac atc tca aag caa gct tgg gaa aca ctt tgg ccc ctt   1474
Tyr Thr Thr Asn Ile Ser Lys Gln Ala Trp Glu Thr Leu Trp Pro Leu
    420                 425                 430

gaa agg aaa aga cag aga gca ttc ttt ctc ttt gga tta gca ctg att   1522
Glu Arg Lys Arg Gln Arg Ala Phe Phe Leu Phe Gly Leu Ala Leu Ile
435                 440                 445                 450

gtc cag atg gat att gag ggg acc cgc aca ttc ttc cgg act ttc ttc   1570
Val Gln Met Asp Ile Glu Gly Thr Arg Thr Phe Phe Arg Thr Phe Phe
                455                 460                 465

cgc ttg ccc aca tgg atg tgg tgg ggg ttt ctt gga tct tcg tta tca   1618
Arg Leu Pro Thr Trp Met Trp Trp Gly Phe Leu Gly Ser Ser Leu Ser
            470                 475                 480

tca act gac ttg ata ata ttt gcg ttt tac atg ttt atc ata gca ccg   1666
Ser Thr Asp Leu Ile Ile Phe Ala Phe Tyr Met Phe Ile Ile Ala Pro
        485                 490                 495

cat agc ctg aga atg ggt ctg gtt aga cat ttg ctt tct gac ccg aca   1714
His Ser Leu Arg Met Gly Leu Val Arg His Leu Leu Ser Asp Pro Thr
    500                 505                 510

gga gga aca atg tta aaa gcg tat ctc acg ata taaataactc tagtcgcgat 1767
Gly Gly Thr Met Leu Lys Ala Tyr Leu Thr Ile
515                 520                 525

cagtttagat tataggcaca tcttgcatat atatatgtat aaaccttatg tgtgctgtat 1827

ccttacatca acacagtcat taattgtatt tcttggggta atgctgatga agtattttct 1887

ggaaaaaaaa aaaaaaaaaa aaaaaaaaa                                    1916
```

<210> 12
<211> 525
<212> PRT
<213> Tagetes erecta

<400> 12

Met Ser Met Arg Ala Gly His Met Thr Ala Thr Met Ala Ala Phe Thr
1 5 10 15

Cys Pro Arg Phe Met Thr Ser Ile Arg Tyr Thr Lys Gln Ile Lys Cys
20 25 30

Asn Ala Ala Lys Ser Gln Leu Val Val Lys Gln Glu Ile Glu Glu Glu
35 40 45

Glu Asp Tyr Val Lys Ala Gly Gly Ser Glu Leu Leu Phe Val Gln Met
50 55 60

Gln Gln Asn Lys Ser Met Asp Ala Gln Ser Ser Leu Ser Gln Lys Leu
65 70 75 80

Pro Arg Val Pro Ile Gly Gly Gly Gly Asp Ser Asn Cys Ile Leu Asp
85 90 95

Leu Val Val Ile Gly Cys Gly Pro Ala Gly Leu Ala Leu Ala Gly Glu
100 105 110

Ser Ala Lys Leu Gly Leu Asn Val Ala Leu Ile Gly Pro Asp Leu Pro
115 120 125

Phe Thr Asn Asn Tyr Gly Val Trp Glu Asp Glu Phe Ile Gly Leu Gly
130 135 140

Leu Glu Gly Cys Ile Glu His Val Trp Arg Asp Thr Val Val Tyr Leu
145 150 155 160

Asp Asp Asn Asp Pro Ile Leu Ile Gly Arg Ala Tyr Gly Arg Val Ser
165 170 175

Arg Asp Leu Leu His Glu Glu Leu Leu Thr Arg Cys Met Glu Ser Gly
180 185 190

Val Ser Tyr Leu Ser Ser Lys Val Glu Arg Ile Thr Glu Ala Pro Asn
195 200 205

Gly Leu Ser Leu Ile Glu Cys Glu Gly Asn Ile Thr Ile Pro Cys Arg
210 215 220

Leu Ala Thr Val Ala Ser Gly Ala Ala Ser Gly Lys Leu Leu Gln Tyr
225 230 235 240

Glu Leu Gly Gly Pro Arg Val Cys Val Gln Thr Ala Tyr Gly Tyr Glu
245 250 255

Val Glu Val Glu Ser Ile Pro Tyr Asp Pro Ser Leu Met Val Phe Met
260 265 270

Asp Tyr Arg Asp Tyr Thr Lys His Lys Ser Gln Ser Leu Glu Ala Gln
275 280 285

Tyr Pro Thr Phe Leu Tyr Val Met Pro Met Ser Pro Thr Lys Val Phe
290 295 300

Phe Glu Glu Thr Cys Leu Ala Ser Lys Glu Ala Met Pro Phe Glu Leu
305 310 315 320

Leu Lys Thr Lys Leu Met Ser Arg Leu Lys Thr Met Gly Ile Arg Ile
325 330 335

Thr Lys Thr Tyr Glu Glu Tyr Leu Val Ala Cys Gln Tyr Leu Glu Glu
340 345 350

Trp Ser Tyr Ile Pro Val Gly Gly Ser Leu Pro Asn Thr Glu Gln Lys
355 360 365

```
Asn Leu Ala Phe Gly Ala Ala Ala Ser Met Val His Pro Ala Thr Gly
    370                 375                 380
Tyr Ser Val Val Arg Ser Leu Ser Glu Ala Pro Asn Tyr Ala Ala Val
385                 390                 395                 400
Ile Ala Lys Ile Leu Gly Lys Gly Asn Ser Lys Gln Met Leu Asp Leu
                405                 410                 415
Gly Arg Tyr Thr Thr Asn Ile Ser Lys Gln Ala Trp Glu Thr Leu Trp
            420                 425                 430
Pro Leu Glu Arg Lys Arg Gln Arg Ala Phe Phe Leu Phe Gly Leu Ala
        435                 440                 445
Leu Ile Val Gln Met Asp Ile Glu Gly Thr Arg Thr Phe Phe Arg Thr
    450                 455                 460
Phe Phe Arg Leu Pro Thr Trp Met Trp Trp Gly Phe Leu Gly Ser Ser
465                 470                 475                 480
Leu Ser Ser Thr Asp Leu Ile Ile Phe Ala Phe Tyr Met Phe Ile Ile
                485                 490                 495
Ala Pro His Ser Leu Arg Met Gly Leu Val Arg His Leu Leu Ser Asp
            500                 505                 510
Pro Thr Gly Gly Thr Met Leu Lys Ala Tyr Leu Thr Ile
        515                 520                 525
```

<210> 13
<211> 1818
<212> DNA
<213> Arabidopsis thaliana

<220>
<221> CDS
<222> (104)..(1675)
<223> coding for epsilon-cyclase

<400> 13

```
aggaaataat tagattcctc tttctgcttg ctataccttg atagaacaat ataacaatgg 60
tgtaagtctt ctcgctgtat tcgaaattat ttggaggagg aaa atg gag tgt gtt   115
                                               Met Glu Cys Val
                                                 1
ggg gct agg aat ttc gca gca atg gcg gtt tca aca ttt ccg tca tgg   163
Gly Ala Arg Asn Phe Ala Ala Met Ala Val Ser Thr Phe Pro Ser Trp
  5                  10                  15                  20
agt tgt cga agg aaa ttt cca gtg gtt aag aga tac agc tat agg aat   211
Ser Cys Arg Arg Lys Phe Pro Val Val Lys Arg Tyr Ser Tyr Arg Asn
                 25                  30                  35
att cgt ttc ggt ttg tgt agt gtc aga gct agc ggc ggc gga agt tcc   259
Ile Arg Phe Gly Leu Cys Ser Val Arg Ala Ser Gly Gly Gly Ser Ser
             40                  45                  50
ggt agt gag agt tgt gta gcg gtg aga gaa gat ttc gct gac gaa gaa   307
Gly Ser Glu Ser Cys Val Ala Val Arg Glu Asp Phe Ala Asp Glu Glu
         55                  60                  65
gat ttt gtg aaa gct ggt ggt tct gag att cta ttt gtt caa atg cag   355
Asp Phe Val Lys Ala Gly Gly Ser Glu Ile Leu Phe Val Gln Met Gln
     70                  75                  80
```

```
cag aac aaa gat atg gat gaa cag tct aag ctt gtt gat aag ttg cct   403
Gln Asn Lys Asp Met Asp Glu Gln Ser Lys Leu Val Asp Lys Leu Pro
 85                  90                  95                 100

cct ata tca att ggt gat ggt gct ttg gat cta gtg gtt att ggt tgt   451
Pro Ile Ser Ile Gly Asp Gly Ala Leu Asp Leu Val Val Ile Gly Cys
                105                 110                 115

ggt cct gct ggt tta gcc ttg gct gca gaa tca gct aag ctt gga tta   499
Gly Pro Ala Gly Leu Ala Leu Ala Ala Glu Ser Ala Lys Leu Gly Leu
            120                 125                 130

aaa gtt gga ctc att ggt cca gat ctt cct ttt act aac aat tac ggt   547
Lys Val Gly Leu Ile Gly Pro Asp Leu Pro Phe Thr Asn Asn Tyr Gly
        135                 140                 145

gtt tgg gaa gat gaa ttc aat gat ctt ggg ctg caa aaa tgt att gag   595
Val Trp Glu Asp Glu Phe Asn Asp Leu Gly Leu Gln Lys Cys Ile Glu
    150                 155                 160

cat gtt tgg aga gag act att gtg tat ctg gat gat gac aag cct att   643
His Val Trp Arg Glu Thr Ile Val Tyr Leu Asp Asp Asp Lys Pro Ile
165                 170                 175                 180

acc att ggc cgt gct tat gga aga gtt agt cga cgt ttg ctc cat gag   691
Thr Ile Gly Arg Ala Tyr Gly Arg Val Ser Arg Arg Leu Leu His Glu
                185                 190                 195

gag ctt ttg agg agg tgt gtc gag tca ggt gtc tcg tac ctt agc tcg   739
Glu Leu Leu Arg Arg Cys Val Glu Ser Gly Val Ser Tyr Leu Ser Ser
            200                 205                 210

aaa gtt gac agc ata aca gaa gct tct gat ggc ctt aga ctt gtt gct   787
Lys Val Asp Ser Ile Thr Glu Ala Ser Asp Gly Leu Arg Leu Val Ala
        215                 220                 225

tgt gac gac aat aac gtc att ccc tgc agg ctt gcc act gtt gct tct   835
Cys Asp Asp Asn Asn Val Ile Pro Cys Arg Leu Ala Thr Val Ala Ser
    230                 235                 240

gga gca gct tcg gga aag ctc ttg caa tac gaa gtt ggt gga cct aga   883
Gly Ala Ala Ser Gly Lys Leu Leu Gln Tyr Glu Val Gly Gly Pro Arg
245                 250                 255                 260

gtc tgt gtg caa act gca tac ggc gtg gag gtt gag gtg gaa aat agt   931
Val Cys Val Gln Thr Ala Tyr Gly Val Glu Val Glu Val Glu Asn Ser
                265                 270                 275

cca tat gat cca gat caa atg gtt ttc atg gat tac aga gat tat act   979
Pro Tyr Asp Pro Asp Gln Met Val Phe Met Asp Tyr Arg Asp Tyr Thr
            280                 285                 290

aac gag aaa gtt cgg agc tta gaa gct gag tat cca acg ttt ctg tac  1027
Asn Glu Lys Val Arg Ser Leu Glu Ala Glu Tyr Pro Thr Phe Leu Tyr
        295                 300                 305

gcc atg cct atg aca aag tca aga ctc ttc ttc gag gag aca tgt ttg  1075
Ala Met Pro Met Thr Lys Ser Arg Leu Phe Phe Glu Glu Thr Cys Leu
    310                 315                 320

gcc tca aaa gat gtc atg ccc ttt gat ttg cta aaa acg aag ctc atg  1123
Ala Ser Lys Asp Val Met Pro Phe Asp Leu Leu Lys Thr Lys Leu Met
325                 330                 335                 340
```

```
tta aga tta gat aca ctc gga att cga att cta aag act tac gaa gag   1171
Leu Arg Leu Asp Thr Leu Gly Ile Arg Ile Leu Lys Thr Tyr Glu Glu
                345                 350                 355

gag tgg tcc tat atc cca gtt ggt ggt tcc ttg cca aac acc gaa caa   1219
Glu Trp Ser Tyr Ile Pro Val Gly Gly Ser Leu Pro Asn Thr Glu Gln
            360                 365                 370

aag aat ctc gcc ttt ggt gct gcc gct agc atg gta cat ccc gca aca   1267
Lys Asn Leu Ala Phe Gly Ala Ala Ala Ser Met Val His Pro Ala Thr
        375                 380                 385

ggc tat tca gtt gtg aga tct ttg tct gaa gct cca aaa tat gca tca   1315
Gly Tyr Ser Val Val Arg Ser Leu Ser Glu Ala Pro Lys Tyr Ala Ser
    390                 395                 400

gtc atc gca gag ata cta aga gaa gag act acc aaa cag atc aac agt   1363
Val Ile Ala Glu Ile Leu Arg Glu Glu Thr Thr Lys Gln Ile Asn Ser
405                 410                 415                 420

aat att tca aga caa gct tgg gat act tta tgg cca cca gaa agg aaa   1411
Asn Ile Ser Arg Gln Ala Trp Asp Thr Leu Trp Pro Pro Glu Arg Lys
                425                 430                 435

aga cag aga gca ttc ttt ctc ttt ggt ctt gca ctc ata gtt caa ttc   1459
Arg Gln Arg Ala Phe Phe Leu Phe Gly Leu Ala Leu Ile Val Gln Phe
            440                 445                 450

gat acc gaa ggc att aga agc ttc ttc cgt act ttc ttc cgc ctt cca   1507
Asp Thr Glu Gly Ile Arg Ser Phe Phe Arg Thr Phe Phe Arg Leu Pro
        455                 460                 465

aaa tgg atg tgg caa ggg ttt cta gga tca aca tta aca tca gga gat   1555
Lys Trp Met Trp Gln Gly Phe Leu Gly Ser Thr Leu Thr Ser Gly Asp
    470                 475                 480

ctc gtt ctc ttt gct tta tac atg ttc gtc att tca cca aac aat ttg   1603
Leu Val Leu Phe Ala Leu Tyr Met Phe Val Ile Ser Pro Asn Asn Leu
485                 490                 495                 500

aga aaa ggt ctc atc aat cat ctc atc tct gat cca acc gga gca acc   1651
Arg Lys Gly Leu Ile Asn His Leu Ile Ser Asp Pro Thr Gly Ala Thr
                505                 510                 515

atg ata aaa acc tat ctc aaa gta tgatttactt atcaactctt aggtttgtgt  1705
Met Ile Lys Thr Tyr Leu Lys Val
            520

atatatatgt tgatttatct gaataatcga tcaaagaatg gtatgtgggt tactaggaag 1765

ttggaaacaa acatgtatag aatctaagga gtgatcgaaa tggagatgga aac         1818
```

<210> 14
<211> 524
<212> PRT
<213> Arabidopsis thaliana

<400> 14

```
Met Glu Cys Val Gly Ala Arg Asn Phe Ala Ala Met Ala Val Ser Thr
  1               5                  10                  15

Phe Pro Ser Trp Ser Cys Arg Arg Lys Phe Pro Val Val Lys Arg Tyr
            20                  25                  30

Ser Tyr Arg Asn Ile Arg Phe Gly Leu Cys Ser Val Arg Ala Ser Gly
        35                  40                  45
```

Gly Gly Ser Ser Gly Ser Glu Ser Cys Val Ala Val Arg Glu Asp Phe
50 55 60

Ala Asp Glu Glu Asp Phe Val Lys Ala Gly Gly Ser Glu Ile Leu Phe
65 70 75 80

Val Gln Met Gln Gln Asn Lys Asp Met Asp Glu Gln Ser Lys Leu Val
85 90 95

Asp Lys Leu Pro Pro Ile Ser Ile Gly Asp Gly Ala Leu Asp Leu Val
100 105 110

Val Ile Gly Cys Gly Pro Ala Gly Leu Ala Leu Ala Ala Glu Ser Ala
115 120 125

Lys Leu Gly Leu Lys Val Gly Leu Ile Gly Pro Asp Leu Pro Phe Thr
130 135 140

Asn Asn Tyr Gly Val Trp Glu Asp Glu Phe Asn Asp Leu Gly Leu Gln
145 150 155 160

Lys Cys Ile Glu His Val Trp Arg Glu Thr Ile Val Tyr Leu Asp Asp
165 170 175

Asp Lys Pro Ile Thr Ile Gly Arg Ala Tyr Gly Arg Val Ser Arg Arg
180 185 190

Leu Leu His Glu Glu Leu Leu Arg Arg Cys Val Glu Ser Gly Val Ser
195 200 205

Tyr Leu Ser Ser Lys Val Asp Ser Ile Thr Glu Ala Ser Asp Gly Leu
210 215 220

Arg Leu Val Ala Cys Asp Asp Asn Asn Val Ile Pro Cys Arg Leu Ala
225 230 235 240

Thr Val Ala Ser Gly Ala Ala Ser Gly Lys Leu Leu Gln Tyr Glu Val
245 250 255

Gly Gly Pro Arg Val Cys Val Gln Thr Ala Tyr Gly Val Glu Val Glu
260 265 270

Val Glu Asn Ser Pro Tyr Asp Pro Asp Gln Met Val Phe Met Asp Tyr
275 280 285

Arg Asp Tyr Thr Asn Glu Lys Val Arg Ser Leu Glu Ala Glu Tyr Pro
290 295 300

Thr Phe Leu Tyr Ala Met Pro Met Thr Lys Ser Arg Leu Phe Phe Glu
305 310 315 320

Glu Thr Cys Leu Ala Ser Lys Asp Val Met Pro Phe Asp Leu Leu Lys
325 330 335

Thr Lys Leu Met Leu Arg Leu Asp Thr Leu Gly Ile Arg Ile Leu Lys
340 345 350

Thr Tyr Glu Glu Glu Trp Ser Tyr Ile Pro Val Gly Gly Ser Leu Pro
355 360 365

Asn Thr Glu Gln Lys Asn Leu Ala Phe Gly Ala Ala Ala Ser Met Val
370 375 380

His Pro Ala Thr Gly Tyr Ser Val Val Arg Ser Leu Ser Glu Ala Pro
385 390 395 400

Lys Tyr Ala Ser Val Ile Ala Glu Ile Leu Arg Glu Glu Thr Thr Lys
405 410 415

Gln Ile Asn Ser Asn Ile Ser Arg Gln Ala Trp Asp Thr Leu Trp Pro
420 425 430

```
Pro Glu Arg Lys Arg Gln Arg Ala Phe Phe Leu Phe Gly Leu Ala Leu
        435                 440                 445
Ile Val Gln Phe Asp Thr Glu Gly Ile Arg Ser Phe Phe Arg Thr Phe
    450                 455                 460
Phe Arg Leu Pro Lys Trp Met Trp Gln Gly Phe Leu Gly Ser Thr Leu
465                 470                 475                 480
Thr Ser Gly Asp Leu Val Leu Phe Ala Leu Tyr Met Phe Val Ile Ser
                485                 490                 495
Pro Asn Asn Leu Arg Lys Gly Leu Ile Asn His Leu Ile Ser Asp Pro
            500                 505                 510
Thr Gly Ala Thr Met Ile Lys Thr Tyr Leu Lys Val
        515                 520
```

<210> 15
<211> 1623
<212> DNA
<213> Oryza sativa

<220>
<221> CDS
<222> (1)..(1620)
<223> coding for epsilon-cyclase

<400> 15

```
atg gag ttc tcc ggc ggc gcg acc gtg tcg gcg ccg ttc ggt tgc tgc   48
Met Glu Phe Ser Gly Gly Ala Thr Val Ser Ala Pro Phe Gly Cys Cys
  1               5                  10                  15

cgt gcg gcg tgg ggc gcc gcg gcg gcg ggt gcg ggg gcg gag gga agg   96
Arg Ala Ala Trp Gly Ala Ala Ala Ala Gly Ala Gly Ala Glu Gly Arg
             20                  25                  30

agc agg agg gtt gtg ccg cgc gcg gtg gag ccg cgg cgg cgc ggg cgg   144
Ser Arg Arg Val Val Pro Arg Ala Val Glu Pro Arg Arg Arg Gly Arg
         35                  40                  45

tgg atg gtg agg tgc gtg gcg acg gag aag cac aag gac gcg gcg gcg   192
Trp Met Val Arg Cys Val Ala Thr Glu Lys His Lys Asp Ala Ala Ala
     50                  55                  60

cgg cgc ggc ggc gtg gag gtg gag ttc gcc gac gag gag gac tac gtc   240
Arg Arg Gly Gly Val Glu Val Glu Phe Ala Asp Glu Glu Asp Tyr Val
 65                  70                  75                  80

aag ggc ggc ggc ggc gag ctt ctc tac gtg caa atg cag gcg tcc aag   288
Lys Gly Gly Gly Gly Glu Leu Leu Tyr Val Gln Met Gln Ala Ser Lys
                 85                  90                  95

tcc atg gac agc cag tcc aag atc tcc tcc aag ctg ctg ccc ata ccc   336
Ser Met Asp Ser Gln Ser Lys Ile Ser Ser Lys Leu Leu Pro Ile Pro
            100                 105                 110

gat gaa aat tca gtt ctt gat ttg gtc atc att ggc tgc ggt cca gct   384
Asp Glu Asn Ser Val Leu Asp Leu Val Ile Ile Gly Cys Gly Pro Ala
        115                 120                 125

ggc tta tcc cta gca gca gag tca gct aag aaa ggg ctc aat gtt ggt   432
Gly Leu Ser Leu Ala Ala Glu Ser Ala Lys Lys Gly Leu Asn Val Gly
    130                 135                 140
```

```
ctc att ggc cct gat ctt cca ttc acg aac aac tac ggt gtg tgg gag   480
Leu Ile Gly Pro Asp Leu Pro Phe Thr Asn Asn Tyr Gly Val Trp Glu
145                 150                 155                 160

gat gaa ttc aaa gac ctg ggc ctg gag agc tgc att gaa cat gtc tgg   528
Asp Glu Phe Lys Asp Leu Gly Leu Glu Ser Cys Ile Glu His Val Trp
                165                 170                 175

aag gat act atc gtg tac cta gac ggt aac aag cca ata atg att ggc   576
Lys Asp Thr Ile Val Tyr Leu Asp Gly Asn Lys Pro Ile Met Ile Gly
            180                 185                 190

cgt gcg tat ggc agg gtt cac agg gac ttg ctg cac gag gag ttg ctg   624
Arg Ala Tyr Gly Arg Val His Arg Asp Leu Leu His Glu Glu Leu Leu
        195                 200                 205

aga cga tgc tat gac gct ggc gtc aca tac ctc agc tcg aag gtg gac   672
Arg Arg Cys Tyr Asp Ala Gly Val Thr Tyr Leu Ser Ser Lys Val Asp
    210                 215                 220

aag atc atg gaa tct cct gat gga cat cgg gta gtc tgt tgt gaa ggg   720
Lys Ile Met Glu Ser Pro Asp Gly His Arg Val Val Cys Cys Glu Gly
225                 230                 235                 240

gat cgt gag gta ctt tgc agg ctt gcc att gtt gca tct ggg gca gca   768
Asp Arg Glu Val Leu Cys Arg Leu Ala Ile Val Ala Ser Gly Ala Ala
                245                 250                 255

tct ggt agg ctt cta gag tac gag gtt ggt ggt ccg cgt gtt tgt gtg   816
Ser Gly Arg Leu Leu Glu Tyr Glu Val Gly Gly Pro Arg Val Cys Val
            260                 265                 270

cag act gca tat ggt gtc gaa gtc gag gtg gaa aac aat cca tat gat   864
Gln Thr Ala Tyr Gly Val Glu Val Glu Val Glu Asn Asn Pro Tyr Asp
        275                 280                 285

ccc agc tta atg gtt ttc atg gac tac aga gat tgc ttc aaa gac aaa   912
Pro Ser Leu Met Val Phe Met Asp Tyr Arg Asp Cys Phe Lys Asp Lys
    290                 295                 300

ttc tca cat cct gag caa gga aat cca acg ttc ctc tat gcc atg ccc   960
Phe Ser His Pro Glu Gln Gly Asn Pro Thr Phe Leu Tyr Ala Met Pro
305                 310                 315                 320

atg tca tcc aca cga att ttc ttt gag gaa aca tgc cta gct tct aaa  1008
Met Ser Ser Thr Arg Ile Phe Phe Glu Glu Thr Cys Leu Ala Ser Lys
                325                 330                 335

gaa gca atg ccc ttt gac ctc ctt aaa aag cgg ttg atg tct cgg ttg  1056
Glu Ala Met Pro Phe Asp Leu Leu Lys Lys Arg Leu Met Ser Arg Leu
            340                 345                 350

gat gca atg gga gtt cat att cga aaa gta tac gag gag gaa tgg tcc  1104
Asp Ala Met Gly Val His Ile Arg Lys Val Tyr Glu Glu Glu Trp Ser
        355                 360                 365

tac att cct gtt gga ggg tcc tta cca aat aca gac cag aaa aat ctc  1152
Tyr Ile Pro Val Gly Gly Ser Leu Pro Asn Thr Asp Gln Lys Asn Leu
    370                 375                 380

gca ttt ggt gcg gca gca agt atg gtg cat cct gca acc gga tac tcg  1200
Ala Phe Gly Ala Ala Ala Ser Met Val His Pro Ala Thr Gly Tyr Ser
385                 390                 395                 400
```

```
gtg gtt aga tca ttg tct gaa gct cca aga tat gca tct gtg ata tct   1248
Val Val Arg Ser Leu Ser Glu Ala Pro Arg Tyr Ala Ser Val Ile Ser
                405                 410                 415

gat atc ttg aga aac cgt gtc tac cct gga gaa tat ttg cct gga acc   1296
Asp Ile Leu Arg Asn Arg Val Tyr Pro Gly Glu Tyr Leu Pro Gly Thr
            420                 425                 430

tct caa agt tcc agt cca tca atg ctt gca tgg aga aca tta tgg ccc   1344
Ser Gln Ser Ser Ser Pro Ser Met Leu Ala Trp Arg Thr Leu Trp Pro
        435                 440                 445

caa gaa cgg aaa cgt caa cga tca ttc ttc ctt ttt ggg ctg gct ttg   1392
Gln Glu Arg Lys Arg Gln Arg Ser Phe Phe Leu Phe Gly Leu Ala Leu
    450                 455                 460

ata atc caa ctg aat aac gaa ggc att cag aca ttc ttt gaa acc ttt   1440
Ile Ile Gln Leu Asn Asn Glu Gly Ile Gln Thr Phe Phe Glu Thr Phe
465                 470                 475                 480

ttc cgg ttg ccc aaa tgg atg tgg cga gga ttc ctt ggt tcg acg ctt   1488
Phe Arg Leu Pro Lys Trp Met Trp Arg Gly Phe Leu Gly Ser Thr Leu
                485                 490                 495

tct tca gtg gat ctc ata ctc ttt gca ttc tac atg ttc aca att gcg   1536
Ser Ser Val Asp Leu Ile Leu Phe Ala Phe Tyr Met Phe Thr Ile Ala
            500                 505                 510

ccg aac caa atg cga atg aac ctt gtc aga cat ctc ctc tct gat ccg   1584
Pro Asn Gln Met Arg Met Asn Leu Val Arg His Leu Leu Ser Asp Pro
        515                 520                 525

acc ggc tca acg atg atc aag acc tac ctg acc ttg taa               1623
Thr Gly Ser Thr Met Ile Lys Thr Tyr Leu Thr Leu
    530                 535                 540
```

<210> 16
<211> 540
<212> PRT
<213> Oryza sativa

<400> 16

```
Met Glu Phe Ser Gly Gly Ala Thr Val Ser Ala Pro Phe Gly Cys Cys
  1               5                  10                  15

Arg Ala Ala Trp Gly Ala Ala Ala Ala Gly Ala Gly Ala Glu Gly Arg
             20                  25                  30

Ser Arg Arg Val Val Pro Arg Ala Val Glu Pro Arg Arg Arg Gly Arg
         35                  40                  45

Trp Met Val Arg Cys Val Ala Thr Glu Lys His Lys Asp Ala Ala Ala
     50                  55                  60

Arg Arg Gly Gly Val Glu Val Glu Phe Ala Asp Glu Glu Asp Tyr Val
 65                  70                  75                  80

Lys Gly Gly Gly Gly Glu Leu Leu Tyr Val Gln Met Gln Ala Ser Lys
                 85                  90                  95

Ser Met Asp Ser Gln Ser Lys Ile Ser Ser Lys Leu Leu Pro Ile Pro
            100                 105                 110

Asp Glu Asn Ser Val Leu Asp Leu Val Ile Ile Gly Cys Gly Pro Ala
        115                 120                 125

Gly Leu Ser Leu Ala Ala Glu Ser Ala Lys Lys Gly Leu Asn Val Gly
    130                 135                 140
```

Leu Ile Gly Pro Asp Leu Pro Phe Thr Asn Asn Tyr Gly Val Trp Glu
145 150 155 160

Asp Glu Phe Lys Asp Leu Gly Leu Glu Ser Cys Ile Glu His Val Trp
165 170 175

Lys Asp Thr Ile Val Tyr Leu Asp Gly Asn Lys Pro Ile Met Ile Gly
180 185 190

Arg Ala Tyr Gly Arg Val His Arg Asp Leu Leu His Glu Glu Leu Leu
195 200 205

Arg Arg Cys Tyr Asp Ala Gly Val Thr Tyr Leu Ser Ser Lys Val Asp
210 215 220

Lys Ile Met Glu Ser Pro Asp Gly His Arg Val Val Cys Cys Glu Gly
225 230 235 240

Asp Arg Glu Val Leu Cys Arg Leu Ala Ile Val Ala Ser Gly Ala Ala
245 250 255

Ser Gly Arg Leu Leu Glu Tyr Glu Val Gly Gly Pro Arg Val Cys Val
260 265 270

Gln Thr Ala Tyr Gly Val Glu Val Glu Val Glu Asn Asn Pro Tyr Asp
275 280 285

Pro Ser Leu Met Val Phe Met Asp Tyr Arg Asp Cys Phe Lys Asp Lys
290 295 300

Phe Ser His Pro Glu Gln Gly Asn Pro Thr Phe Leu Tyr Ala Met Pro
305 310 315 320

Met Ser Ser Thr Arg Ile Phe Phe Glu Glu Thr Cys Leu Ala Ser Lys
325 330 335

Glu Ala Met Pro Phe Asp Leu Leu Lys Lys Arg Leu Met Ser Arg Leu
340 345 350

Asp Ala Met Gly Val His Ile Arg Lys Val Tyr Glu Glu Glu Trp Ser
355 360 365

Tyr Ile Pro Val Gly Gly Ser Leu Pro Asn Thr Asp Gln Lys Asn Leu
370 375 380

Ala Phe Gly Ala Ala Ala Ser Met Val His Pro Ala Thr Gly Tyr Ser
385 390 395 400

Val Val Arg Ser Leu Ser Glu Ala Pro Arg Tyr Ala Ser Val Ile Ser
405 410 415

Asp Ile Leu Arg Asn Arg Val Tyr Pro Gly Glu Tyr Leu Pro Gly Thr
420 425 430

Ser Gln Ser Ser Ser Pro Ser Met Leu Ala Trp Arg Thr Leu Trp Pro
435 440 445

Gln Glu Arg Lys Arg Gln Arg Ser Phe Phe Leu Phe Gly Leu Ala Leu
450 455 460

Ile Ile Gln Leu Asn Asn Glu Gly Ile Gln Thr Phe Phe Glu Thr Phe
465 470 475 480

Phe Arg Leu Pro Lys Trp Met Trp Arg Gly Phe Leu Gly Ser Thr Leu
485 490 495

Ser Ser Val Asp Leu Ile Leu Phe Ala Phe Tyr Met Phe Thr Ile Ala
500 505 510

Pro Asn Gln Met Arg Met Asn Leu Val Arg His Leu Leu Ser Asp Pro
515 520 525

Thr Gly Ser Thr Met Ile Lys Thr Tyr Leu Thr Leu
530 535 540

<210> 17
<211> 10
<212> PRT
<213> Artificial sequence

<220>
<223> Description of the artificial sequence: protein motif for epsilon-cyclase

<220>
<221> VARIANT
<222> (2)
<223> G/C variation

<220>
<221> VARIANT
<222> (8)
<223> A/S variation

<220>
<221> VARIANT
<222> (9)
<223> V/L variation

<400> 17
Gly Gly Gly Pro Ala Gly Leu Ala Val Ala
1 5 10

<210> 18
<211> 8
<212> PRT
<213> Artificial sequence

<220>
<223> Description of the artificial sequence: protein motif for epsilon-cyclase

<220>
<221> VARIANT
<222> (1)
<223> L/I variation

<220>
<221> VARIANT
<222> (2)
<223> N/G/S variation

<220>
<221> VARIANT
<222> (7)
<223> K/R variation

<220>
<221> VARIANT
<222> (8)
<223> V/L variation

<400> 18
Leu Asn Arg Xaa Tyr Gly Lys Val
1 5

<210> 19
<211> 8
<212> PRT
<213> Artificial sequence

<220>
<223> Description of the artificial sequence: protein motif for epsilon-cyclase

<220>
<221> VARIANT
<222> (6)
<223> Y/W variation

<400> 19
Met Val Phe Met Asp Tyr Arg Asp
1 5

<210> 20
<211> 8
<212> PRT
<213> Artificial sequence

<220>
<223> Description of the artificial sequence: protein motif for epsilon-cyclase

<220>
<221> VARIANT
<222> (6)
<223> A/V variation

<220>
<221> VARIANT
<222> (8)
<223> P/A variation

<400> 20
Pro Thr Phe Leu Tyr Ala Met Pro
1 5

<210> 21
<211> 14
<212> PRT
<213> Artificial sequence

<220>
<223> Description of the artificial sequence: protein motif for epsilon-cyclase

<220>
<221> VARIANT
<222> (7)
<223> S/A-variation

<220>
<221> VARIANT
<222> (11)
<223> M/S variation

<220>
<221> VARIANT
<222> (13)
<223> A/V variation

<400> 21
Ala Xaa Met Val His Pro Ser Thr Gly Tyr Met Val Ala Arg
1 5 10

<210> 22
<211> 13
<212> PRT
<213> Artificial sequence

<220>
<223> Description of the artificial sequence: protein motif for epsilon-cyclase

<220>
<221> VARIANT
<222> (7)
<223> R/K variation

<400> 22
Leu Trp Pro Xaa Glu Arg Arg Arg Gln Arg Xaa Phe Phe
1 5 10

<210> 23
<211> 1780
<212> DNA
<213> Lactuca sativa

<220>
<221> CDS
<222> (77)..(1675)
<223> coding for epsilon-cyclase

<400> 23
gaaacaaatg acgtgaaagt tcttcaaaat tgaattaatt gtaatcctga aaacttgatt 60

tgtgatagaa gaatca atg gag tgc ttt gga gct cga aac atg acg gca aca 112
Met Glu Cys Phe Gly Ala Arg Asn Met Thr Ala Thr
1 5 10

atg gcg gtt ttt acg tgc cct aga ttc acg gac tgt aat atc agg cac 160
Met Ala Val Phe Thr Cys Pro Arg Phe Thr Asp Cys Asn Ile Arg His
15 20 25

aaa ttt tcg tta ctg aaa caa cga aga ttt act aat tta tca gca tcg 208
Lys Phe Ser Leu Leu Lys Gln Arg Arg Phe Thr Asn Leu Ser Ala Ser
30 35 40

tct tcg ttg cgt caa att aag tgc agc gct aaa agc gac cgt tgt gta 256
Ser Ser Leu Arg Gln Ile Lys Cys Ser Ala Lys Ser Asp Arg Cys Val
45 50 55 60

gtg gat aaa caa ggg att tcc gta gca gac gaa gaa gat tat gtg aag 304
Val Asp Lys Gln Gly Ile Ser Val Ala Asp Glu Glu Asp Tyr Val Lys
65 70 75

gcc ggt gga tcg gag ctg ttt ttt gtt caa atg cag cgg act aag tcc 352
Ala Gly Gly Ser Glu Leu Phe Phe Val Gln Met Gln Arg Thr Lys Ser
80 85 90

atg gaa agc cag tct aaa ctt tcc gaa aag cta gca cag ata cca att 400
Met Glu Ser Gln Ser Lys Leu Ser Glu Lys Leu Ala Gln Ile Pro Ile
95 100 105

gga aat tgc ata ctt gat ctg gtt gta atc ggt tgt ggc cct gct ggc 448
Gly Asn Cys Ile Leu Asp Leu Val Val Ile Gly Cys Gly Pro Ala Gly
110 115 120

```
ctt gct ctt gct gca gag tca gcc aaa cta ggg ttg aac gtt gga ctc   496
Leu Ala Leu Ala Ala Glu Ser Ala Lys Leu Gly Leu Asn Val Gly Leu
125                 130                 135                 140

att ggc cct gat ctt cct ttt aca aac aat tat ggt gtt tgg cag gat   544
Ile Gly Pro Asp Leu Pro Phe Thr Asn Asn Tyr Gly Val Trp Gln Asp
                145                 150                 155

gaa ttt ata ggt ctt gga ctt gaa gga tgc att gaa cat tct tgg aaa   592
Glu Phe Ile Gly Leu Gly Leu Glu Gly Cys Ile Glu His Ser Trp Lys
            160                 165                 170

gat act ctt gta tac ctt gat gat gct gat ccc atc cgc ata ggt cgt   640
Asp Thr Leu Val Tyr Leu Asp Asp Ala Asp Pro Ile Arg Ile Gly Arg
        175                 180                 185

gca tat ggc aga gtt cat cgt gat tta ctt cat gaa gag ttg tta aga   688
Ala Tyr Gly Arg Val His Arg Asp Leu Leu His Glu Glu Leu Leu Arg
    190                 195                 200

agg tgt gtg gaa tca ggt gtt tca tat cta agc tcc aaa gta gaa aga   736
Arg Cys Val Glu Ser Gly Val Ser Tyr Leu Ser Ser Lys Val Glu Arg
205                 210                 215                 220

atc act gaa gct cca aat ggc tat agt ctc att gaa tgt gaa ggc aat   784
Ile Thr Glu Ala Pro Asn Gly Tyr Ser Leu Ile Glu Cys Glu Gly Asn
                225                 230                 235

atc acc att cca tgc agg ctt gct act gtt gca tca ggg gca gct tca   832
Ile Thr Ile Pro Cys Arg Leu Ala Thr Val Ala Ser Gly Ala Ala Ser
            240                 245                 250

ggg aaa ttt ctg gag tat gaa ctt ggg ggt ccc cgt gtt tgt gtc caa   880
Gly Lys Phe Leu Glu Tyr Glu Leu Gly Gly Pro Arg Val Cys Val Gln
        255                 260                 265

aca gct tat ggt ata gag gtt gag gtt gaa aac aac ccc tat gat cca   928
Thr Ala Tyr Gly Ile Glu Val Glu Val Glu Asn Asn Pro Tyr Asp Pro
    270                 275                 280

gat cta atg gtg ttc atg gat tat aga gac ttc tca aaa cat aaa ccg   976
Asp Leu Met Val Phe Met Asp Tyr Arg Asp Phe Ser Lys His Lys Pro
285                 290                 295                 300

gaa tct tta gaa gca aaa tat ccg act ttc ctc tat gtc atg gcc atg   1024
Glu Ser Leu Glu Ala Lys Tyr Pro Thr Phe Leu Tyr Val Met Ala Met
                305                 310                 315

tct cca aca aaa ata ttc ttc gag gaa act tgt tta gct tca aga gaa   1072
Ser Pro Thr Lys Ile Phe Phe Glu Glu Thr Cys Leu Ala Ser Arg Glu
            320                 325                 330

gcc atg cct ttc aat ctt cta aag tcc aaa ctc atg tca cga tta aag   1120
Ala Met Pro Phe Asn Leu Leu Lys Ser Lys Leu Met Ser Arg Leu Lys
        335                 340                 345

gca atg ggt atc cga ata aca aga acg tac gaa gag gaa tgg tcg tat   1168
Ala Met Gly Ile Arg Ile Thr Arg Thr Tyr Glu Glu Glu Trp Ser Tyr
    350                 355                 360

atc ccc gta ggt gga tcg tta cct aat aca gaa caa aag aat ctc gca   1216
Ile Pro Val Gly Gly Ser Leu Pro Asn Thr Glu Gln Lys Asn Leu Ala
365                 370                 375                 380
```

```
ttt ggt gct gca gct agt atg gtg cac cct gcc aca ggg tat tca gtt   1264
Phe Gly Ala Ala Ala Ser Met Val His Pro Ala Thr Gly Tyr Ser Val
                385                 390                 395

gtt cga tct ttg tca gaa gct cct aat tat gca gca gtc att gct aag   1312
Val Arg Ser Leu Ser Glu Ala Pro Asn Tyr Ala Ala Val Ile Ala Lys
            400                 405                 410

att tta aga caa gat caa tct aaa gag atg att tct ctt gga aaa tac   1360
Ile Leu Arg Gln Asp Gln Ser Lys Glu Met Ile Ser Leu Gly Lys Tyr
        415                 420                 425

act aac att tca aaa caa gca tgg gaa aca ttg tgg cca ctt gaa agg   1408
Thr Asn Ile Ser Lys Gln Ala Trp Glu Thr Leu Trp Pro Leu Glu Arg
    430                 435                 440

aaa aga cag cga gcc ttc ttt cta ttc gga cta tca cac atc gtg cta   1456
Lys Arg Gln Arg Ala Phe Phe Leu Phe Gly Leu Ser His Ile Val Leu
445                 450                 455                 460

atg gat cta gag gga aca cgt aca ttt ttc cgt act ttc ttt cgt ttg   1504
Met Asp Leu Glu Gly Thr Arg Thr Phe Phe Arg Thr Phe Phe Arg Leu
                465                 470                 475

ccc aaa tgg atg tgg tgg gga ttt ttg ggg tct tct tta tct tca acg   1552
Pro Lys Trp Met Trp Trp Gly Phe Leu Gly Ser Ser Leu Ser Ser Thr
            480                 485                 490

gat ttg ata ata ttt gcg ctt tat atg ttt gtg ata gca cct cac agc   1600
Asp Leu Ile Ile Phe Ala Leu Tyr Met Phe Val Ile Ala Pro His Ser
        495                 500                 505

ttg aga atg gaa ctg gtt aga cat cta ctt tct gat ccg aca ggg gca   1648
Leu Arg Met Glu Leu Val Arg His Leu Leu Ser Asp Pro Thr Gly Ala
    510                 515                 520

act atg gta aaa gca tat ctc act ata tagatttaga ttatataaat         1695
Thr Met Val Lys Ala Tyr Leu Thr Ile
525                 530

aatacccata tcttgcatat atataagcct tatttatttc ttttgtatcc ttacaacaac 1755

atactcgtta attatatgtt tttta                                       1780
```

<210> 24
<211> 533
<212> PRT
<213> Lactuca sativa

<400> 24

```
Met Glu Cys Phe Gly Ala Arg Asn Met Thr Ala Thr Met Ala Val Phe
  1               5                  10                  15

Thr Cys Pro Arg Phe Thr Asp Cys Asn Ile Arg His Lys Phe Ser Leu
            20                  25                  30

Leu Lys Gln Arg Arg Phe Thr Asn Leu Ser Ala Ser Ser Ser Leu Arg
        35                  40                  45

Gln Ile Lys Cys Ser Ala Lys Ser Asp Arg Cys Val Val Asp Lys Gln
    50                  55                  60

Gly Ile Ser Val Ala Asp Glu Glu Asp Tyr Val Lys Ala Gly Gly Ser
65                  70                  75                  80

Glu Leu Phe Phe Val Gln Met Gln Arg Thr Lys Ser Met Glu Ser Gln
                85                  90                  95
```

Ser Lys Leu Ser Glu Lys Leu Ala Gln Ile Pro Ile Gly Asn Cys Ile
100 105 110

Leu Asp Leu Val Val Ile Gly Cys Gly Pro Ala Gly Leu Ala Leu Ala
115 120 125

Ala Glu Ser Ala Lys Leu Gly Leu Asn Val Gly Leu Ile Gly Pro Asp
130 135 140

Leu Pro Phe Thr Asn Asn Tyr Gly Val Trp Gln Asp Glu Phe Ile Gly
145 150 155 160

Leu Gly Leu Glu Gly Cys Ile Glu His Ser Trp Lys Asp Thr Leu Val
165 170 175

Tyr Leu Asp Asp Ala Asp Pro Ile Arg Ile Gly Arg Ala Tyr Gly Arg
180 185 190

Val His Arg Asp Leu Leu His Glu Glu Leu Leu Arg Arg Cys Val Glu
195 200 205

Ser Gly Val Ser Tyr Leu Ser Ser Lys Val Glu Arg Ile Thr Glu Ala
210 215 220

Pro Asn Gly Tyr Ser Leu Ile Glu Cys Glu Gly Asn Ile Thr Ile Pro
225 230 235 240

Cys Arg Leu Ala Thr Val Ala Ser Gly Ala Ala Ser Gly Lys Phe Leu
245 250 255

Glu Tyr Glu Leu Gly Gly Pro Arg Val Cys Val Gln Thr Ala Tyr Gly
260 265 270

Ile Glu Val Glu Val Glu Asn Asn Pro Tyr Asp Pro Asp Leu Met Val
275 280 285

Phe Met Asp Tyr Arg Asp Phe Ser Lys His Lys Pro Glu Ser Leu Glu
290 295 300

Ala Lys Tyr Pro Thr Phe Leu Tyr Val Met Ala Met Ser Pro Thr Lys
305 310 315 320

Ile Phe Phe Glu Glu Thr Cys Leu Ala Ser Arg Glu Ala Met Pro Phe
325 330 335

Asn Leu Leu Lys Ser Lys Leu Met Ser Arg Leu Lys Ala Met Gly Ile
340 345 350

Arg Ile Thr Arg Thr Tyr Glu Glu Glu Trp Ser Tyr Ile Pro Val Gly
355 360 365

Gly Ser Leu Pro Asn Thr Glu Gln Lys Asn Leu Ala Phe Gly Ala Ala
370 375 380

Ala Ser Met Val His Pro Ala Thr Gly Tyr Ser Val Val Arg Ser Leu
385 390 395 400

Ser Glu Ala Pro Asn Tyr Ala Ala Val Ile Ala Lys Ile Leu Arg Gln
405 410 415

Asp Gln Ser Lys Glu Met Ile Ser Leu Gly Lys Tyr Thr Asn Ile Ser
420 425 430

Lys Gln Ala Trp Glu Thr Leu Trp Pro Leu Glu Arg Lys Arg Gln Arg
435 440 445

Ala Phe Phe Leu Phe Gly Leu Ser His Ile Val Leu Met Asp Leu Glu
450 455 460

Gly Thr Arg Thr Phe Phe Arg Thr Phe Phe Arg Leu Pro Lys Trp Met
465 470 475 480

```
Trp Trp Gly Phe Leu Gly Ser Ser Leu Ser Ser Thr Asp Leu Ile Ile
                485                 490                 495
Phe Ala Leu Tyr Met Phe Val Ile Ala Pro His Ser Leu Arg Met Glu
            500                 505                 510
Leu Val Arg His Leu Leu Ser Asp Pro Thr Gly Ala Thr Met Val Lys
        515                 520                 525
Ala Tyr Leu Thr Ile
    530
```

<210> 25
<211> 1848
<212> DNA
<213> Adonis palaestina

<220>
<221> CDS
<222> (116)..(1702)
<223> coding for epsilon cyclase

<400> 25

```
gagagaaaaa gagtgttata ttaatgttac tgtcgcattc ttgcaacaca tattcagact 60
ccatttttctt gttttctctt caaaacaaca aactaatgtg acggagtatc tagct atg  118
                                                            Met
                                                             1
gaa cta ctt ggt gtt cgc aac ctc atc tct tct tgc cct gtc tgg act   166
Glu Leu Leu Gly Val Arg Asn Leu Ile Ser Ser Cys Pro Val Trp Thr
                5                  10                  15
ttt gga aca aga aac ctt agt agt tca aaa cta gct tat aac ata cat   214
Phe Gly Thr Arg Asn Leu Ser Ser Ser Lys Leu Ala Tyr Asn Ile His
            20                  25                  30
cga tat ggt tct tct tgt aga gta gat ttt caa gtg agg gct gat ggt   262
Arg Tyr Gly Ser Ser Cys Arg Val Asp Phe Gln Val Arg Ala Asp Gly
        35                  40                  45
gga agc ggg agt aga act tct gtt gct tat aaa gag ggt ttt gtg gac   310
Gly Ser Gly Ser Arg Thr Ser Val Ala Tyr Lys Glu Gly Phe Val Asp
    50                  55                  60                  65
gag gag gat ttt atc aaa gct ggt ggt tct gag ctt ttg ttt gtc caa   358
Glu Glu Asp Phe Ile Lys Ala Gly Gly Ser Glu Leu Leu Phe Val Gln
                70                  75                  80
atg cag caa aca aag tct atg gag aaa cag gcc aag ctc gcc gat aag   406
Met Gln Gln Thr Lys Ser Met Glu Lys Gln Ala Lys Leu Ala Asp Lys
            85                  90                  95
ttg cca cca ata cct ttc gga gaa tct gtg atg gac ttg gtt gta ata   454
Leu Pro Pro Ile Pro Phe Gly Glu Ser Val Met Asp Leu Val Val Ile
        100                 105                 110
ggt tgt gga cct gct ggt ctt tca ctg gct gca gaa gct gct aag cta   502
Gly Cys Gly Pro Ala Gly Leu Ser Leu Ala Ala Glu Ala Ala Lys Leu
    115                 120                 125
ggc ttg aaa gtt ggc ctt att ggt cct gat ctt cct ttt aca aat aat   550
Gly Leu Lys Val Gly Leu Ile Gly Pro Asp Leu Pro Phe Thr Asn Asn
130                 135                 140                 145
```

```
tat ggt gtg tgg gaa gac gag ttc aaa gat ctt gga ctt gaa cgt tgt   598
Tyr Gly Val Trp Glu Asp Glu Phe Lys Asp Leu Gly Leu Glu Arg Cys
                150                 155                 160

atc gag cat gct tgg aag gac acc atc gta tat ctt gac aat gat gct   646
Ile Glu His Ala Trp Lys Asp Thr Ile Val Tyr Leu Asp Asn Asp Ala
            165                 170                 175

cct gtc ctt att ggt cgt gca tat gga cga gtt agc cgg cat ttg ctg   694
Pro Val Leu Ile Gly Arg Ala Tyr Gly Arg Val Ser Arg His Leu Leu
        180                 185                 190

cat gaa gag ttg ctg aaa agg tgt gtc gag tca ggt gta tca tat ctg   742
His Glu Glu Leu Leu Lys Arg Cys Val Glu Ser Gly Val Ser Tyr Leu
    195                 200                 205

aat tct aaa gtg gaa agg atc act gaa gct ggt gat ggc cat agt ctt   790
Asn Ser Lys Val Glu Arg Ile Thr Glu Ala Gly Asp Gly His Ser Leu
210                 215                 220                 225

gta gtt tgt gaa aac gac atc ttt atc cct tgc agg ctt gct act gtt   838
Val Val Cys Glu Asn Asp Ile Phe Ile Pro Cys Arg Leu Ala Thr Val
                230                 235                 240

gca tct gga gca gct tca ggg aaa ctt ttg gag tat gaa gta ggt ggc   886
Ala Ser Gly Ala Ala Ser Gly Lys Leu Leu Glu Tyr Glu Val Gly Gly
            245                 250                 255

cct cgt gtt tgt gtc caa act gct tat ggt gtg gag gtt gag gtg gag   934
Pro Arg Val Cys Val Gln Thr Ala Tyr Gly Val Glu Val Glu Val Glu
        260                 265                 270

aac aat cca tac gat ccc aac tta atg gta ttt atg gac tac aga gac   982
Asn Asn Pro Tyr Asp Pro Asn Leu Met Val Phe Met Asp Tyr Arg Asp
    275                 280                 285

tat atg caa cag aaa tta cag tgc tcg gaa gaa gaa tat cca aca ttt  1030
Tyr Met Gln Gln Lys Leu Gln Cys Ser Glu Glu Glu Tyr Pro Thr Phe
290                 295                 300                 305

ctc tat gtc atg ccc atg tcg cca aca aga ctt ttt ttt gag gaa acc  1078
Leu Tyr Val Met Pro Met Ser Pro Thr Arg Leu Phe Phe Glu Glu Thr
                310                 315                 320

tgt ttg gcc tca aaa gat gcc atg cct ttc gat cta ctg aag aga aaa  1126
Cys Leu Ala Ser Lys Asp Ala Met Pro Phe Asp Leu Leu Lys Arg Lys
            325                 330                 335

cta atg tca cga ttg aag act ctg ggt atc caa gtt aca aaa att tat  1174
Leu Met Ser Arg Leu Lys Thr Leu Gly Ile Gln Val Thr Lys Ile Tyr
        340                 345                 350

gaa gag gaa tgg tct tat att cct gtt ggg ggt tct tta cca aac aca  1222
Glu Glu Glu Trp Ser Tyr Ile Pro Val Gly Gly Ser Leu Pro Asn Thr
    355                 360                 365

gag caa aag aac cta gca ttt ggt gct gca gca agc atg gtg cat cca  1270
Glu Gln Lys Asn Leu Ala Phe Gly Ala Ala Ala Ser Met Val His Pro
370                 375                 380                 385

gca aca ggc tat tcg gtt gta cga tca cta tca gaa gct cca aaa tat  1318
Ala Thr Gly Tyr Ser Val Val Arg Ser Leu Ser Glu Ala Pro Lys Tyr
                390                 395                 400
```

```
gct tct gta att gca aag att ttg aag caa gat aac tct gca tat gtg   1366
Ala Ser Val Ile Ala Lys Ile Leu Lys Gln Asp Asn Ser Ala Tyr Val
            405                 410                 415

gtt tct gga caa agc agt gca gta aac att tca atg caa gca tgg agc   1414
Val Ser Gly Gln Ser Ser Ala Val Asn Ile Ser Met Gln Ala Trp Ser
        420                 425                 430

agt ctt tgg cca aag gag cga aaa cgt caa aga gca ttc ttt ctt ttc   1462
Ser Leu Trp Pro Lys Glu Arg Lys Arg Gln Arg Ala Phe Phe Leu Phe
    435                 440                 445

ggg tta gag ctt att gtg cag cta gat att gaa gca acc aga acg ttc   1510
Gly Leu Glu Leu Ile Val Gln Leu Asp Ile Glu Ala Thr Arg Thr Phe
450                 455                 460                 465

ttt aga acc ttc ttc cgc ttg cca act tgg atg tgg tgg ggt ttc ctt   1558
Phe Arg Thr Phe Phe Arg Leu Pro Thr Trp Met Trp Trp Gly Phe Leu
                470                 475                 480

ggg tct tca cta tca tct ttc gat ctt gta ttg ttt tcc atg tac atg   1606
Gly Ser Ser Leu Ser Ser Phe Asp Leu Val Leu Phe Ser Met Tyr Met
            485                 490                 495

ttt gtt ttg gcc ccg aac agc atg agg atg tca ctt gtg aga cat ttg   1654
Phe Val Leu Ala Pro Asn Ser Met Arg Met Ser Leu Val Arg His Leu
        500                 505                 510

ctt tca gat cct tct ggt gca gtt atg gtt aaa gct tac ctc gaa agg   1702
Leu Ser Asp Pro Ser Gly Ala Val Met Val Lys Ala Tyr Leu Glu Arg
    515                 520                 525

taatctgttt tatgaaacta tagtgtctca ttaaataaat gaggatcctt cgtatatgta 1762

tatgatcatc tctatgtata tcctatattc taatctcata aagtaatcga aaattcattg 1822

atagaaaaaa aaaaaaaaaa aaaaaa                                      1848
```

<210> 26
<211> 529
<212> PRT
<213> Adonis palaestina

<400> 26

```
Met Glu Leu Leu Gly Val Arg Asn Leu Ile Ser Ser Cys Pro Val Trp
  1               5                  10                  15

Thr Phe Gly Thr Arg Asn Leu Ser Ser Ser Lys Leu Ala Tyr Asn Ile
             20                  25                  30

His Arg Tyr Gly Ser Ser Cys Arg Val Asp Phe Gln Val Arg Ala Asp
         35                  40                  45

Gly Gly Ser Gly Ser Arg Thr Ser Val Ala Tyr Lys Glu Gly Phe Val
     50                  55                  60

Asp Glu Glu Asp Phe Ile Lys Ala Gly Gly Ser Glu Leu Leu Phe Val
 65                  70                  75                  80

Gln Met Gln Gln Thr Lys Ser Met Glu Lys Gln Ala Lys Leu Ala Asp
                 85                  90                  95

Lys Leu Pro Pro Ile Pro Phe Gly Glu Ser Val Met Asp Leu Val Val
            100                 105                 110

Ile Gly Cys Gly Pro Ala Gly Leu Ser Leu Ala Ala Glu Ala Ala Lys
        115                 120                 125
```

Leu Gly Leu Lys Val Gly Leu Ile Gly Pro Asp Leu Pro Phe Thr Asn
130 135 140
Asn Tyr Gly Val Trp Glu Asp Glu Phe Lys Asp Leu Gly Leu Glu Arg
145 150 155 160
Cys Ile Glu His Ala Trp Lys Asp Thr Ile Val Tyr Leu Asp Asn Asp
165 170 175
Ala Pro Val Leu Ile Gly Arg Ala Tyr Gly Arg Val Ser Arg His Leu
180 185 190
Leu His Glu Glu Leu Leu Lys Arg Cys Val Glu Ser Gly Val Ser Tyr
195 200 205
Leu Asn Ser Lys Val Glu Arg Ile Thr Glu Ala Gly Asp Gly His Ser
210 215 220
Leu Val Val Cys Glu Asn Asp Ile Phe Ile Pro Cys Arg Leu Ala Thr
225 230 235 240
Val Ala Ser Gly Ala Ala Ser Gly Lys Leu Leu Glu Tyr Glu Val Gly
245 250 255
Gly Pro Arg Val Cys Val Gln Thr Ala Tyr Gly Val Glu Val Glu Val
260 265 270
Glu Asn Asn Pro Tyr Asp Pro Asn Leu Met Val Phe Met Asp Tyr Arg
275 280 285
Asp Tyr Met Gln Gln Lys Leu Gln Cys Ser Glu Glu Glu Tyr Pro Thr
290 295 300
Phe Leu Tyr Val Met Pro Met Ser Pro Thr Arg Leu Phe Phe Glu Glu
305 310 315 320
Thr Cys Leu Ala Ser Lys Asp Ala Met Pro Phe Asp Leu Leu Lys Arg
325 330 335
Lys Leu Met Ser Arg Leu Lys Thr Leu Gly Ile Gln Val Thr Lys Ile
340 345 350
Tyr Glu Glu Glu Trp Ser Tyr Ile Pro Val Gly Gly Ser Leu Pro Asn
355 360 365
Thr Glu Gln Lys Asn Leu Ala Phe Gly Ala Ala Ala Ser Met Val His
370 375 380
Pro Ala Thr Gly Tyr Ser Val Val Arg Ser Leu Ser Glu Ala Pro Lys
385 390 395 400
Tyr Ala Ser Val Ile Ala Lys Ile Leu Lys Gln Asp Asn Ser Ala Tyr
405 410 415
Val Val Ser Gly Gln Ser Ser Ala Val Asn Ile Ser Met Gln Ala Trp
420 425 430
Ser Ser Leu Trp Pro Lys Glu Arg Lys Arg Gln Arg Ala Phe Phe Leu
435 440 445
Phe Gly Leu Glu Leu Ile Val Gln Leu Asp Ile Glu Ala Thr Arg Thr
450 455 460
Phe Phe Arg Thr Phe Phe Arg Leu Pro Thr Trp Met Trp Trp Gly Phe
465 470 475 480
Leu Gly Ser Ser Leu Ser Ser Phe Asp Leu Val Leu Phe Ser Met Tyr
485 490 495
Met Phe Val Leu Ala Pro Asn Ser Met Arg Met Ser Leu Val Arg His
500 505 510

Leu Leu Ser Asp Pro Ser Gly Ala Val Met Val Lys Ala Tyr Leu Glu
515 520 525

Arg

<210> 27
<211> 1898
<212> DNA
<213> Adonis palaestina

<220>
<221> CDS
<222> (113)..(1699)
<223> coding for epsilon-cyclase

<400> 27
aaaggagtgt tctattaatg ttactgtcgc attcttgcaa cacttatatt caaactccat 60

tttcttcttt tctcttcaaa acaacaaact aatgtgagca gagtatctgg ct atg gaa 118
Met Glu
1

cta ctt ggt gtt cgc aac ctc atc tct tct tgc cct gtg tgg act ttt 166
Leu Leu Gly Val Arg Asn Leu Ile Ser Ser Cys Pro Val Trp Thr Phe
5 10 15

gga aca aga aac ctt agt agt tca aaa cta gct tat aac ata cat cga 214
Gly Thr Arg Asn Leu Ser Ser Ser Lys Leu Ala Tyr Asn Ile His Arg
20 25 30

tat ggt tct tct tgt aga gta gat ttt caa gtg aga gct gat ggt gga 262
Tyr Gly Ser Ser Cys Arg Val Asp Phe Gln Val Arg Ala Asp Gly Gly
35 40 45 50

agc ggg agt aga agt tct gtt gct tat aaa gag ggt ttt gtg gat gaa 310
Ser Gly Ser Arg Ser Ser Val Ala Tyr Lys Glu Gly Phe Val Asp Glu
55 60 65

gag gat ttt atc aaa gct ggt ggt tct gag ctt ttg ttt gtc caa atg 358
Glu Asp Phe Ile Lys Ala Gly Gly Ser Glu Leu Leu Phe Val Gln Met
70 75 80

cag caa aca aag tct atg gag aaa cag gcc aag ctc gcc gat aag ttg 406
Gln Gln Thr Lys Ser Met Glu Lys Gln Ala Lys Leu Ala Asp Lys Leu
85 90 95

cca cca ata cct ttt gga gaa tcc gtg atg gac ttg gtt gta ata ggt 454
Pro Pro Ile Pro Phe Gly Glu Ser Val Met Asp Leu Val Val Ile Gly
100 105 110

tgt gga cct gct ggt ctt tca ctg gct gca gaa gct gct aag cta ggg 502
Cys Gly Pro Ala Gly Leu Ser Leu Ala Ala Glu Ala Ala Lys Leu Gly
115 120 125 130

ttg aaa gtt ggc ctt att ggt cct gat ctt cct ttt aca aat aat tat 550
Leu Lys Val Gly Leu Ile Gly Pro Asp Leu Pro Phe Thr Asn Asn Tyr
135 140 145

ggt gtg tgg gaa gac gag ttc aaa gat ctt gga ctt gaa cgt tgt atc 598
Gly Val Trp Glu Asp Glu Phe Lys Asp Leu Gly Leu Glu Arg Cys Ile
150 155 160

```
gag cat gct tgg aag gac acc atc gta tat ctt gat aat gat gct cct   646
Glu His Ala Trp Lys Asp Thr Ile Val Tyr Leu Asp Asn Asp Ala Pro
        165                 170                 175

gtc ctt att ggt cgt gca tat gga cga gtt agt cga cat ttg cta cat   694
Val Leu Ile Gly Arg Ala Tyr Gly Arg Val Ser Arg His Leu Leu His
    180                 185                 190

gag gag ttg ctg aaa agg tgt gtg gag tca ggt gta tca tat ctg gat   742
Glu Glu Leu Leu Lys Arg Cys Val Glu Ser Gly Val Ser Tyr Leu Asp
195                 200                 205                 210

tct aaa gtg gaa agg atc act gaa gct ggt gat ggc cat agc ctt gta   790
Ser Lys Val Glu Arg Ile Thr Glu Ala Gly Asp Gly His Ser Leu Val
                215                 220                 225

gtt tgt gaa aat gag atc ttt atc cct tgc agg ctt gct act gtt gca   838
Val Cys Glu Asn Glu Ile Phe Ile Pro Cys Arg Leu Ala Thr Val Ala
            230                 235                 240

tct gga gca gct tca ggg aaa ctt ttg gag tat gaa gta ggt ggc cct   886
Ser Gly Ala Ala Ser Gly Lys Leu Leu Glu Tyr Glu Val Gly Gly Pro
        245                 250                 255

cgt gtt tgt gtc caa acc gct tat ggg gtg gag gtt gag gtg gag aac   934
Arg Val Cys Val Gln Thr Ala Tyr Gly Val Glu Val Glu Val Glu Asn
    260                 265                 270

aat cca tac gat ccc aac tta atg gta ttc atg gac tac aga gac tat   982
Asn Pro Tyr Asp Pro Asn Leu Met Val Phe Met Asp Tyr Arg Asp Tyr
275                 280                 285                 290

atg caa cag aaa tta cag tgc tcg gaa gaa gaa tat cca aca ttt ctc  1030
Met Gln Gln Lys Leu Gln Cys Ser Glu Glu Glu Tyr Pro Thr Phe Leu
                295                 300                 305

tat gtc atg ccc atg tcg cca aca aga ctt ttt ttt gag gaa acc tgt  1078
Tyr Val Met Pro Met Ser Pro Thr Arg Leu Phe Phe Glu Glu Thr Cys
            310                 315                 320

ttg gcc tca aaa gat gcc atg cca ttc gat cta ctg aag aga aaa ctg  1126
Leu Ala Ser Lys Asp Ala Met Pro Phe Asp Leu Leu Lys Arg Lys Leu
        325                 330                 335

atg tca cga ttg aag act ctg ggt atc caa gtt aca aaa gtt tat gaa  1174
Met Ser Arg Leu Lys Thr Leu Gly Ile Gln Val Thr Lys Val Tyr Glu
    340                 345                 350

gag gaa tgg tca tat att cct gtt ggt ggt tct tta cca aac aca gag  1222
Glu Glu Trp Ser Tyr Ile Pro Val Gly Gly Ser Leu Pro Asn Thr Glu
355                 360                 365                 370

caa aag aac cta gca ttt ggt gct gca gca agc atg gtg cat cca gca  1270
Gln Lys Asn Leu Ala Phe Gly Ala Ala Ala Ser Met Val His Pro Ala
                375                 380                 385

aca ggc tat tcg gtt gta cgg tca ctg tca gaa gct cca aaa tat gct  1318
Thr Gly Tyr Ser Val Val Arg Ser Leu Ser Glu Ala Pro Lys Tyr Ala
            390                 395                 400

tct gta att gca aag att ttg aag caa gat aac tct gcg tat gtg gtt  1366
Ser Val Ile Ala Lys Ile Leu Lys Gln Asp Asn Ser Ala Tyr Val Val
        405                 410                 415
```

```
tct gga caa agt agt gca gta aac att tca atg caa gca tgg agc agt   1414
Ser Gly Gln Ser Ser Ala Val Asn Ile Ser Met Gln Ala Trp Ser Ser
    420                 425                 430

ctt tgg cca aag gag cga aaa cgt caa aga gca ttc ttt ctt ttt gga   1462
Leu Trp Pro Lys Glu Arg Lys Arg Gln Arg Ala Phe Phe Leu Phe Gly
435                 440                 445                 450

tta gag ctt att gtg cag cta gat att gaa gca acc aga aca ttc ttt   1510
Leu Glu Leu Ile Val Gln Leu Asp Ile Glu Ala Thr Arg Thr Phe Phe
                455                 460                 465

aga acc ttc ttc cgc ttg cca act tgg atg tgg tgg ggt ttc ctt ggg   1558
Arg Thr Phe Phe Arg Leu Pro Thr Trp Met Trp Trp Gly Phe Leu Gly
            470                 475                 480

tct tca cta tca tct ttc gat ctc gtc ttg ttt tcc atg tac atg ttt   1606
Ser Ser Leu Ser Ser Phe Asp Leu Val Leu Phe Ser Met Tyr Met Phe
        485                 490                 495

gtt ttg gcg cca aac agc atg agg atg tca ctt gtg aga cat ttg ctt   1654
Val Leu Ala Pro Asn Ser Met Arg Met Ser Leu Val Arg His Leu Leu
    500                 505                 510

tca gat cct tct ggt gca gtt atg gta aga gct tac ctc gaa agg       1699
Ser Asp Pro Ser Gly Ala Val Met Val Arg Ala Tyr Leu Glu Arg
515                 520                 525

tagtctcatc tattattaaa ctctagtgtt tcaccaaata aatgaggatc cttcgaatgt 1759

gtatatgatc atctctatgt atatcctgta ctctaatctc ataaagtaaa tgccgggttt 1819

gatattgttg tgtcaaaccg gccaatgata taaagtaaat ttattgatac aaaagtagtt 1879

tttttcctta aaaaaaaaa                                            1898
```

<210> 28
<211> 529
<212> PRT
<213> Adonis palaestina

<400> 28

```
Met Glu Leu Leu Gly Val Arg Asn Leu Ile Ser Ser Cys Pro Val Trp
  1               5                  10                  15

Thr Phe Gly Thr Arg Asn Leu Ser Ser Ser Lys Leu Ala Tyr Asn Ile
            20                  25                  30

His Arg Tyr Gly Ser Ser Cys Arg Val Asp Phe Gln Val Arg Ala Asp
        35                  40                  45

Gly Gly Ser Gly Ser Arg Ser Ser Val Ala Tyr Lys Glu Gly Phe Val
    50                  55                  60

Asp Glu Glu Asp Phe Ile Lys Ala Gly Gly Ser Glu Leu Leu Phe Val
65                  70                  75                  80

Gln Met Gln Gln Thr Lys Ser Met Glu Lys Gln Ala Lys Leu Ala Asp
                85                  90                  95

Lys Leu Pro Pro Ile Pro Phe Gly Glu Ser Val Met Asp Leu Val Val
            100                 105                 110

Ile Gly Cys Gly Pro Ala Gly Leu Ser Leu Ala Ala Glu Ala Ala Lys
        115                 120                 125

Leu Gly Leu Lys Val Gly Leu Ile Gly Pro Asp Leu Pro Phe Thr Asn
    130                 135                 140
```

```
Asn Tyr Gly Val Trp Glu Asp Glu Phe Lys Asp Leu Gly Leu Glu Arg
145                 150                 155                 160

Cys Ile Glu His Ala Trp Lys Asp Thr Ile Val Tyr Leu Asp Asn Asp
                165                 170                 175

Ala Pro Val Leu Ile Gly Arg Ala Tyr Gly Arg Val Ser Arg His Leu
            180                 185                 190

Leu His Glu Glu Leu Leu Lys Arg Cys Val Glu Ser Gly Val Ser Tyr
        195                 200                 205

Leu Asp Ser Lys Val Glu Arg Ile Thr Glu Ala Gly Asp Gly His Ser
    210                 215                 220

Leu Val Val Cys Glu Asn Glu Ile Phe Ile Pro Cys Arg Leu Ala Thr
225                 230                 235                 240

Val Ala Ser Gly Ala Ala Ser Gly Lys Leu Leu Glu Tyr Glu Val Gly
                245                 250                 255

Gly Pro Arg Val Cys Val Gln Thr Ala Tyr Gly Val Glu Val Glu Val
            260                 265                 270

Glu Asn Asn Pro Tyr Asp Pro Asn Leu Met Val Phe Met Asp Tyr Arg
        275                 280                 285

Asp Tyr Met Gln Gln Lys Leu Gln Cys Ser Glu Glu Glu Tyr Pro Thr
    290                 295                 300

Phe Leu Tyr Val Met Pro Met Ser Pro Thr Arg Leu Phe Phe Glu Glu
305                 310                 315                 320

Thr Cys Leu Ala Ser Lys Asp Ala Met Pro Phe Asp Leu Leu Lys Arg
                325                 330                 335

Lys Leu Met Ser Arg Leu Lys Thr Leu Gly Ile Gln Val Thr Lys Val
            340                 345                 350

Tyr Glu Glu Glu Trp Ser Tyr Ile Pro Val Gly Gly Ser Leu Pro Asn
        355                 360                 365

Thr Glu Gln Lys Asn Leu Ala Phe Gly Ala Ala Ala Ser Met Val His
    370                 375                 380

Pro Ala Thr Gly Tyr Ser Val Val Arg Ser Leu Ser Glu Ala Pro Lys
385                 390                 395                 400

Tyr Ala Ser Val Ile Ala Lys Ile Leu Lys Gln Asp Asn Ser Ala Tyr
                405                 410                 415

Val Val Ser Gly Gln Ser Ser Ala Val Asn Ile Ser Met Gln Ala Trp
            420                 425                 430

Ser Ser Leu Trp Pro Lys Glu Arg Lys Arg Gln Arg Ala Phe Phe Leu
        435                 440                 445

Phe Gly Leu Glu Leu Ile Val Gln Leu Asp Ile Glu Ala Thr Arg Thr
    450                 455                 460

Phe Phe Arg Thr Phe Phe Arg Leu Pro Thr Trp Met Trp Trp Gly Phe
465                 470                 475                 480

Leu Gly Ser Ser Leu Ser Ser Phe Asp Leu Val Leu Phe Ser Met Tyr
                485                 490                 495

Met Phe Val Leu Ala Pro Asn Ser Met Arg Met Ser Leu Val Arg His
            500                 505                 510

Leu Leu Ser Asp Pro Ser Gly Ala Val Met Val Arg Ala Tyr Leu Glu
        515                 520                 525
```

Arg

<210> 29
<211> 1661
<212> DNA
<213> Arabidopsis thaliana

<220>
<221> CDS
<222> (2)..(1501)
<223> coding for epsilon-cyclase

<400> 29

```
g atg gat act ctg ttg aaa aca ccc aac aag ctc gat ttt ttc atc cct 49
  Met Asp Thr Leu Leu Lys Thr Pro Asn Lys Leu Asp Phe Phe Ile Pro
    1               5                  10                  15

cag ttt cat ggg ttt gag aga tta tgc agt aac aat cca tac cct tca   97
Gln Phe His Gly Phe Glu Arg Leu Cys Ser Asn Asn Pro Tyr Pro Ser
            20                  25                  30

agg gtt agg ctt ggt gtg aag aaa agg gct atc aaa att gtc tct agt   145
Arg Val Arg Leu Gly Val Lys Lys Arg Ala Ile Lys Ile Val Ser Ser
        35                  40                  45

gta gtg agt ggt agc gct gct ctt ttg gat ctt gtt cct gaa act aag   193
Val Val Ser Gly Ser Ala Ala Leu Leu Asp Leu Val Pro Glu Thr Lys
    50                  55                  60

aag gag aat ctt gac ttt gag ctt cct ttg tac gac act tcc aag agt   241
Lys Glu Asn Leu Asp Phe Glu Leu Pro Leu Tyr Asp Thr Ser Lys Ser
65                  70                  75                  80

caa gtt gtt gat ttg gct att gtt ggt ggt ggt cct gct ggt tta gcc   289
Gln Val Val Asp Leu Ala Ile Val Gly Gly Gly Pro Ala Gly Leu Ala
                85                  90                  95

gtg gct cag cag gtt tct gaa gct gga ctc tct gtt tgt tcc att gat   337
Val Ala Gln Gln Val Ser Glu Ala Gly Leu Ser Val Cys Ser Ile Asp
            100                 105                 110

cct tct cct aag ctc ata tgg cct aac aat tat gga gtt tgg gtt gat   385
Pro Ser Pro Lys Leu Ile Trp Pro Asn Asn Tyr Gly Val Trp Val Asp
        115                 120                 125

gag ttt gag gct atg gat tta cta gac tgc ctg gat acc aca tgg tct   433
Glu Phe Glu Ala Met Asp Leu Leu Asp Cys Leu Asp Thr Thr Trp Ser
    130                 135                 140

ggt gct gtt gtc tat gtc gat gaa ggt gtc aag aag gat ttg agc cgg   481
Gly Ala Val Val Tyr Val Asp Glu Gly Val Lys Lys Asp Leu Ser Arg
145                 150                 155                 160

cct tat ggg aga gtt aac cgg aaa cag ctc aaa tcc aaa atg ctt cag   529
Pro Tyr Gly Arg Val Asn Arg Lys Gln Leu Lys Ser Lys Met Leu Gln
                165                 170                 175

aaa tgt att acc aac ggt gtt aaa ttt cat cag tct aag gtc act aat   577
Lys Cys Ile Thr Asn Gly Val Lys Phe His Gln Ser Lys Val Thr Asn
            180                 185                 190
```

```
gtg gtt cac gag gag gca aac tcc act gtg gtc tgc agt gac ggt gta   625
Val Val His Glu Glu Ala Asn Ser Thr Val Val Cys Ser Asp Gly Val
        195                 200                 205

aag att cag gct tcc gtg gtt ctt gat gcc act ggg ttt tcc cga tgc   673
Lys Ile Gln Ala Ser Val Val Leu Asp Ala Thr Gly Phe Ser Arg Cys
    210                 215                 220

ttg gtt cag tat gac aaa cct tac aac cct ggg tac caa gta gct tac   721
Leu Val Gln Tyr Asp Lys Pro Tyr Asn Pro Gly Tyr Gln Val Ala Tyr
225                 230                 235                 240

ggg att ata gct gaa gtt gat ggt cac cca ttc gat gta gac aaa atg   769
Gly Ile Ile Ala Glu Val Asp Gly His Pro Phe Asp Val Asp Lys Met
                245                 250                 255

gtg ttc atg gat tgg aga gac aaa cat ctg gac tca tat cct gag ctg   817
Val Phe Met Asp Trp Arg Asp Lys His Leu Asp Ser Tyr Pro Glu Leu
            260                 265                 270

aaa gaa cgg aac agc aag atc cca acg ttc ttg tac gct atg cca ttt   865
Lys Glu Arg Asn Ser Lys Ile Pro Thr Phe Leu Tyr Ala Met Pro Phe
        275                 280                 285

tct tcc aac cga ata ttt ctt gaa gaa act tct tta gtt gct aga cct   913
Ser Ser Asn Arg Ile Phe Leu Glu Glu Thr Ser Leu Val Ala Arg Pro
    290                 295                 300

ggt ctg aga atg gaa gat atc caa gaa aga atg gct gct aga ctg aaa   961
Gly Leu Arg Met Glu Asp Ile Gln Glu Arg Met Ala Ala Arg Leu Lys
305                 310                 315                 320

cat ctg ggg atc aat gtg aag agg att gag gaa gac gag cgt tgt gtg  1009
His Leu Gly Ile Asn Val Lys Arg Ile Glu Glu Asp Glu Arg Cys Val
                325                 330                 335

atc ccg atg ggc ggt cct tta cca gtc tta cct caa cgg gtt gtg ggg  1057
Ile Pro Met Gly Gly Pro Leu Pro Val Leu Pro Gln Arg Val Val Gly
            340                 345                 350

att ggt ggg aca gca gga atg gtt cat cct tca act ggt tac atg gtt  1105
Ile Gly Gly Thr Ala Gly Met Val His Pro Ser Thr Gly Tyr Met Val
        355                 360                 365

gct agg act ctt gca gct gca cca ata gtt gca aat gcc att gtg aga  1153
Ala Arg Thr Leu Ala Ala Ala Pro Ile Val Ala Asn Ala Ile Val Arg
    370                 375                 380

tac ctc ggt tca cca agt agt aat agc ctg aga gga gat caa ctc tct  1201
Tyr Leu Gly Ser Pro Ser Ser Asn Ser Leu Arg Gly Asp Gln Leu Ser
385                 390                 395                 400

gct gag gtt tgg aga gac ttg tgg cct atc gaa cgg cgt aga cag agg  1249
Ala Glu Val Trp Arg Asp Leu Trp Pro Ile Glu Arg Arg Arg Gln Arg
                405                 410                 415

gag ttc ttc tgt ttt gga atg gat att ctg ctg aaa ctc gat tta gac  1297
Glu Phe Phe Cys Phe Gly Met Asp Ile Leu Leu Lys Leu Asp Leu Asp
            420                 425                 430

gct act aga agg ttc ttt gat gca ttc ttt gat ctg caa cct cat tac  1345
Ala Thr Arg Arg Phe Phe Asp Ala Phe Phe Asp Leu Gln Pro His Tyr
        435                 440                 445
```

```
tgg cac gga ttc ttg tct tcc agg ctg ttt ctc ccg gaa ctg ttg gtc   1393
Trp His Gly Phe Leu Ser Ser Arg Leu Phe Leu Pro Glu Leu Leu Val
    450                 455                 460

ttc ggg ttg tcg ctc ttc tca cac gct tcc aat acc tca aga ttg gag   1441
Phe Gly Leu Ser Leu Phe Ser His Ala Ser Asn Thr Ser Arg Leu Glu
465                 470                 475                 480

atc atg aca aag ggg act gtt cct ctt gct aag atg atc aac aat ttg   1489
Ile Met Thr Lys Gly Thr Val Pro Leu Ala Lys Met Ile Asn Asn Leu
                485                 490                 495

gta caa gat aga gactaaggac cagaaactta gacatataag tatatctgtt       1541
Val Gln Asp Arg
            500

ctttggttct tgaccagtag tatatccgca ttgcaagtcg ttggataatt gtgtataaac 1601

cacagatcca taacctgaat ccttgtgaaa tcaaattgtt actactagtt cattaaaacc 1661
```

<210> 30
<211> 500
<212> PRT
<213> Arabidopsis thaliana

<400> 30

```
Met Asp Thr Leu Leu Lys Thr Pro Asn Lys Leu Asp Phe Phe Ile Pro
  1               5                  10                  15

Gln Phe His Gly Phe Glu Arg Leu Cys Ser Asn Asn Pro Tyr Pro Ser
            20                  25                  30

Arg Val Arg Leu Gly Val Lys Lys Arg Ala Ile Lys Ile Val Ser Ser
        35                  40                  45

Val Val Ser Gly Ser Ala Ala Leu Leu Asp Leu Val Pro Glu Thr Lys
    50                  55                  60

Lys Glu Asn Leu Asp Phe Glu Leu Pro Leu Tyr Asp Thr Ser Lys Ser
65                  70                  75                  80

Gln Val Val Asp Leu Ala Ile Val Gly Gly Gly Pro Ala Gly Leu Ala
                85                  90                  95

Val Ala Gln Gln Val Ser Glu Ala Gly Leu Ser Val Cys Ser Ile Asp
            100                 105                 110

Pro Ser Pro Lys Leu Ile Trp Pro Asn Asn Tyr Gly Val Trp Val Asp
        115                 120                 125

Glu Phe Glu Ala Met Asp Leu Leu Asp Cys Leu Asp Thr Thr Trp Ser
    130                 135                 140

Gly Ala Val Val Tyr Val Asp Glu Gly Val Lys Lys Asp Leu Ser Arg
145                 150                 155                 160

Pro Tyr Gly Arg Val Asn Arg Lys Gln Leu Lys Ser Lys Met Leu Gln
                165                 170                 175

Lys Cys Ile Thr Asn Gly Val Lys Phe His Gln Ser Lys Val Thr Asn
            180                 185                 190

Val Val His Glu Glu Ala Asn Ser Thr Val Val Cys Ser Asp Gly Val
        195                 200                 205

Lys Ile Gln Ala Ser Val Val Leu Asp Ala Thr Gly Phe Ser Arg Cys
    210                 215                 220

Leu Val Gln Tyr Asp Lys Pro Tyr Asn Pro Gly Tyr Gln Val Ala Tyr
225                 230                 235                 240
```

```
Gly Ile Ile Ala Glu Val Asp Gly His Pro Phe Asp Val Asp Lys Met
                245                 250                 255

Val Phe Met Asp Trp Arg Asp Lys His Leu Asp Ser Tyr Pro Glu Leu
            260                 265                 270

Lys Glu Arg Asn Ser Lys Ile Pro Thr Phe Leu Tyr Ala Met Pro Phe
        275                 280                 285

Ser Ser Asn Arg Ile Phe Leu Glu Glu Thr Ser Leu Val Ala Arg Pro
    290                 295                 300

Gly Leu Arg Met Glu Asp Ile Gln Glu Arg Met Ala Ala Arg Leu Lys
305                 310                 315                 320

His Leu Gly Ile Asn Val Lys Arg Ile Glu Glu Asp Glu Arg Cys Val
                325                 330                 335

Ile Pro Met Gly Gly Pro Leu Pro Val Leu Pro Gln Arg Val Val Gly
            340                 345                 350

Ile Gly Gly Thr Ala Gly Met Val His Pro Ser Thr Gly Tyr Met Val
        355                 360                 365

Ala Arg Thr Leu Ala Ala Ala Pro Ile Val Ala Asn Ala Ile Val Arg
    370                 375                 380

Tyr Leu Gly Ser Pro Ser Ser Asn Ser Leu Arg Gly Asp Gln Leu Ser
385                 390                 395                 400

Ala Glu Val Trp Arg Asp Leu Trp Pro Ile Glu Arg Arg Arg Gln Arg
                405                 410                 415

Glu Phe Phe Cys Phe Gly Met Asp Ile Leu Leu Lys Leu Asp Leu Asp
            420                 425                 430

Ala Thr Arg Arg Phe Phe Asp Ala Phe Phe Asp Leu Gln Pro His Tyr
        435                 440                 445

Trp His Gly Phe Leu Ser Ser Arg Leu Phe Leu Pro Glu Leu Leu Val
    450                 455                 460

Phe Gly Leu Ser Leu Phe Ser His Ala Ser Asn Thr Ser Arg Leu Glu
465                 470                 475                 480

Ile Met Thr Lys Gly Thr Val Pro Leu Ala Lys Met Ile Asn Asn Leu
                485                 490                 495

Val Gln Asp Arg
            500
```

<210> 31
<211> 1550
<212> DNA
<213> Citrus X paradisi

<220>
<221> CDS
<222> (45)..(1355)
<223> coding for epsilon-cyclase

<400> 31
aaataaaaaa cactggcaaa tgctctatta gtacagtgtt atta atg gac atg ttt 56
                                                Met Asp Met Phe
                                                  1

```
ata cta ccg cca ata tca att ggt aat ggt att ttg gat ttg gtg gtg   104
Ile Leu Pro Pro Ile Ser Ile Gly Asn Gly Ile Leu Asp Leu Val Val
  5                  10                  15                  20

att ggt tgt ggc cca gct ggt ctt gct ttg gct gca gaa tca gcg aag   152
Ile Gly Cys Gly Pro Ala Gly Leu Ala Leu Ala Ala Glu Ser Ala Lys
                 25                  30                  35

ttg gga tta aat gtt gga ctt att ggc ccg gat ctc cct ttc aca aac   200
Leu Gly Leu Asn Val Gly Leu Ile Gly Pro Asp Leu Pro Phe Thr Asn
             40                  45                  50

aat tat ggt gtg tgg gaa gat gaa ttt aga gat ctt gga ctt gaa ggg   248
Asn Tyr Gly Val Trp Glu Asp Glu Phe Arg Asp Leu Gly Leu Glu Gly
         55                  60                  65

tgt atc gaa cat gtc tgg aga gac aca gtt gta tat att gat gaa gat   296
Cys Ile Glu His Val Trp Arg Asp Thr Val Val Tyr Ile Asp Glu Asp
     70                  75                  80

gaa ccc atc ttg att ggt cgt gct tat gga cga gtt agt cga cat ttg   344
Glu Pro Ile Leu Ile Gly Arg Ala Tyr Gly Arg Val Ser Arg His Leu
 85                  90                  95                 100

ctt cat gaa gaa tta tta aga agg tgt gtc gag tca ggt gtt tca tat   392
Leu His Glu Glu Leu Leu Arg Arg Cys Val Glu Ser Gly Val Ser Tyr
                105                 110                 115

ctt agc tca aaa gtg gaa agc att acg gaa tct acc agt ggt cat cgt   440
Leu Ser Ser Lys Val Glu Ser Ile Thr Glu Ser Thr Ser Gly His Arg
            120                 125                 130

ctt gta gct tgt gaa cat gat atg att gtc ccc tgc agg ctt gct act   488
Leu Val Ala Cys Glu His Asp Met Ile Val Pro Cys Arg Leu Ala Thr
        135                 140                 145

gtt gct tct gga gca gca tca ggg aag cta ttg gaa tat gag gtg ggg   536
Val Ala Ser Gly Ala Ala Ser Gly Lys Leu Leu Glu Tyr Glu Val Gly
    150                 155                 160

ggt ccc aaa gtt tct gtc caa aca gct tat ggt gtg gag gtt gag gtg   584
Gly Pro Lys Val Ser Val Gln Thr Ala Tyr Gly Val Glu Val Glu Val
165                 170                 175                 180

gaa aat aat cca tat gat cca agc ctt atg gtt ttc atg gac tac aga   632
Glu Asn Asn Pro Tyr Asp Pro Ser Leu Met Val Phe Met Asp Tyr Arg
                185                 190                 195

gac tgt act aag caa gaa gtt cca tct ttt gaa tct gac aat cca aca   680
Asp Cys Thr Lys Gln Glu Val Pro Ser Phe Glu Ser Asp Asn Pro Thr
            200                 205                 210

ttt ctt tat gtc atg ccc atg tct tca aca aga gtt ttc ttt gag gaa   728
Phe Leu Tyr Val Met Pro Met Ser Ser Thr Arg Val Phe Phe Glu Glu
        215                 220                 225

act tgt ttg gca tcg aaa gat ggt tta cgt ttt gac ata ttg aag aaa   776
Thr Cys Leu Ala Ser Lys Asp Gly Leu Arg Phe Asp Ile Leu Lys Lys
    230                 235                 240

aag ctc atg gca agg tta gag aga ttg gga atc cag gtt ttg aaa act   824
Lys Leu Met Ala Arg Leu Glu Arg Leu Gly Ile Gln Val Leu Lys Thr
245                 250                 255                 260
```

```
tat gaa gag gaa tgg tca tat att cca gtt ggt ggt tcc tta cca aat   872
Tyr Glu Glu Glu Trp Ser Tyr Ile Pro Val Gly Gly Ser Leu Pro Asn
                265                 270                 275

aca gaa caa aga aac ctc gca ttt ggt gct gct gct agc atg gtg cat   920
Thr Glu Gln Arg Asn Leu Ala Phe Gly Ala Ala Ala Ser Met Val His
            280                 285                 290

cca gcc act ggc tac tca gta gtc aga tca ctg tca gag gct cca aac   968
Pro Ala Thr Gly Tyr Ser Val Val Arg Ser Leu Ser Glu Ala Pro Asn
        295                 300                 305

tat gct tct gca att gca tat ata ttg aaa cac gat cat tcc aga ggt  1016
Tyr Ala Ser Ala Ile Ala Tyr Ile Leu Lys His Asp His Ser Arg Gly
    310                 315                 320

aga ctt aca cat gaa caa agt aat gag aat atc tca atg caa gct tgg  1064
Arg Leu Thr His Glu Gln Ser Asn Glu Asn Ile Ser Met Gln Ala Trp
325                 330                 335                 340

aat act ctc tgg cca cag gaa agg aag cgc caa aga gct ttt ttc ctc  1112
Asn Thr Leu Trp Pro Gln Glu Arg Lys Arg Gln Arg Ala Phe Phe Leu
                345                 350                 355

ttt gga cta gca ctc att ttg caa ctg gat att gag ggc atc agg aca  1160
Phe Gly Leu Ala Leu Ile Leu Gln Leu Asp Ile Glu Gly Ile Arg Thr
            360                 365                 370

ttc ttt cgc act ttc ttc cga tta ccc aag tgg atg tgg cac ggt ttc  1208
Phe Phe Arg Thr Phe Phe Arg Leu Pro Lys Trp Met Trp His Gly Phe
        375                 380                 385

ctt ggt tct agt ctc tca tca gcc gat ctc att cta ttt gcc ttc tat  1256
Leu Gly Ser Ser Leu Ser Ser Ala Asp Leu Ile Leu Phe Ala Phe Tyr
    390                 395                 400

atg ttt att ata gca cca aat gat ctg aga aag tgc ctt atc aga cat  1304
Met Phe Ile Ile Ala Pro Asn Asp Leu Arg Lys Cys Leu Ile Arg His
405                 410                 415                 420

cta gtt tca gat cca act gga gca act atg gta aga aca tac ctg act  1352
Leu Val Ser Asp Pro Thr Gly Ala Thr Met Val Arg Thr Tyr Leu Thr
                425                 430                 435

tta tagttagttt gtattttcca tatttcagcc cttgtttggt atattttgga       1405
Leu

ttgccatacg tgacacataa tgagcttgta tatatactcc atgtatactg taaactgtta 1465

gtttgacaaa tgaagccctt ttttatttttt aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 1525

aaaaaaaaaa aaaaaaaaaa aaaaa                                      1550
```

<210> 32
<211> 437
<212> PRT
<213> Citrus X paradisi

<400> 32

```
Met Asp Met Phe Ile Leu Pro Pro Ile Ser Ile Gly Asn Gly Ile Leu
  1               5                  10                  15

Asp Leu Val Val Ile Gly Cys Gly Pro Ala Gly Leu Ala Leu Ala Ala
            20                  25                  30

Glu Ser Ala Lys Leu Gly Leu Asn Val Gly Leu Ile Gly Pro Asp Leu
        35                  40                  45
```

```
Pro Phe Thr Asn Asn Tyr Gly Val Trp Glu Asp Glu Phe Arg Asp Leu
     50                  55                  60
Gly Leu Glu Gly Cys Ile Glu His Val Trp Arg Asp Thr Val Val Tyr
 65                  70                  75                  80
Ile Asp Glu Asp Glu Pro Ile Leu Ile Gly Arg Ala Tyr Gly Arg Val
                 85                  90                  95
Ser Arg His Leu Leu His Glu Glu Leu Leu Arg Arg Cys Val Glu Ser
            100                 105                 110
Gly Val Ser Tyr Leu Ser Ser Lys Val Glu Ser Ile Thr Glu Ser Thr
        115                 120                 125
Ser Gly His Arg Leu Val Ala Cys Glu His Asp Met Ile Val Pro Cys
    130                 135                 140
Arg Leu Ala Thr Val Ala Ser Gly Ala Ala Ser Gly Lys Leu Leu Glu
145                 150                 155                 160
Tyr Glu Val Gly Gly Pro Lys Val Ser Val Gln Thr Ala Tyr Gly Val
                165                 170                 175
Glu Val Glu Val Glu Asn Asn Pro Tyr Asp Pro Ser Leu Met Val Phe
            180                 185                 190
Met Asp Tyr Arg Asp Cys Thr Lys Gln Glu Val Pro Ser Phe Glu Ser
        195                 200                 205
Asp Asn Pro Thr Phe Leu Tyr Val Met Pro Met Ser Ser Thr Arg Val
    210                 215                 220
Phe Phe Glu Glu Thr Cys Leu Ala Ser Lys Asp Gly Leu Arg Phe Asp
225                 230                 235                 240
Ile Leu Lys Lys Lys Leu Met Ala Arg Leu Glu Arg Leu Gly Ile Gln
                245                 250                 255
Val Leu Lys Thr Tyr Glu Glu Glu Trp Ser Tyr Ile Pro Val Gly Gly
            260                 265                 270
Ser Leu Pro Asn Thr Glu Gln Arg Asn Leu Ala Phe Gly Ala Ala Ala
        275                 280                 285
Ser Met Val His Pro Ala Thr Gly Tyr Ser Val Val Arg Ser Leu Ser
    290                 295                 300
Glu Ala Pro Asn Tyr Ala Ser Ala Ile Ala Tyr Ile Leu Lys His Asp
305                 310                 315                 320
His Ser Arg Gly Arg Leu Thr His Glu Gln Ser Asn Glu Asn Ile Ser
                325                 330                 335
Met Gln Ala Trp Asn Thr Leu Trp Pro Gln Glu Arg Lys Arg Gln Arg
            340                 345                 350
Ala Phe Phe Leu Phe Gly Leu Ala Leu Ile Leu Gln Leu Asp Ile Glu
        355                 360                 365
Gly Ile Arg Thr Phe Phe Arg Thr Phe Phe Arg Leu Pro Lys Trp Met
    370                 375                 380
Trp His Gly Phe Leu Gly Ser Ser Leu Ser Ser Ala Asp Leu Ile Leu
385                 390                 395                 400
Phe Ala Phe Tyr Met Phe Ile Ile Ala Pro Asn Asp Leu Arg Lys Cys
                405                 410                 415
Leu Ile Arg His Leu Val Ser Asp Pro Thr Gly Ala Thr Met Val Arg
            420                 425                 430
```

Thr Tyr Leu Thr Leu
435

<210> 33
<211> 1830
<212> DNA
<213> Citrus X paradisi

<220>
<221> CDS
<222> (89)..(1660)
<223> coding for epsilon-cyclase

<400> 33
gggttcagtt tgtcgttgag gacaggccac aaacgcaaca caagcttcat ctttaccaaa 60

tttccgtaag caacttctgg gctgaaaa atg ctc cca ttt ctc tcc tct ctg 112
Met Leu Pro Phe Leu Ser Ser Leu
1 5

ctt aat gga gtc acg gat aac cct tgt agg aaa gcc atg gat act tta 160
Leu Asn Gly Val Thr Asp Asn Pro Cys Arg Lys Ala Met Asp Thr Leu
10 15 20

ctc aaa act cat aac aag ctt gaa ttc ttg ccc caa gtt cac ggg gct 208
Leu Lys Thr His Asn Lys Leu Glu Phe Leu Pro Gln Val His Gly Ala
25 30 35 40

ttg gaa aaa tcc agt agt tta agc tca ttg aag att cag aac cag gag 256
Leu Glu Lys Ser Ser Ser Leu Ser Ser Leu Lys Ile Gln Asn Gln Glu
45 50 55

ctt agg ttt ggt ctc aag aag tct cgt caa aag agg aat agg agt tgt 304
Leu Arg Phe Gly Leu Lys Lys Ser Arg Gln Lys Arg Asn Arg Ser Cys
60 65 70

ttc att aag gct agt agt agt gct ctt ttg gag cta gtt cct gaa acc 352
Phe Ile Lys Ala Ser Ser Ser Ala Leu Leu Glu Leu Val Pro Glu Thr
75 80 85

aag aag gaa aat ctt gaa ttt gag ctt ccc atg tat gac cca tca aag 400
Lys Lys Glu Asn Leu Glu Phe Glu Leu Pro Met Tyr Asp Pro Ser Lys
90 95 100

ggc ctt gtt gta gac cta gca gtt gtc ggt ggc ggc ccg gct ggg ctt 448
Gly Leu Val Val Asp Leu Ala Val Val Gly Gly Gly Pro Ala Gly Leu
105 110 115 120

gct gtt gct cag caa gtt tca ggg gcg ggg ctt tcg gtt tgc tcg att 496
Ala Val Ala Gln Gln Val Ser Gly Ala Gly Leu Ser Val Cys Ser Ile
125 130 135

gat cca tct ccc aaa ttg att tgg cca aat aat tat ggt gtt tgg gtg 544
Asp Pro Ser Pro Lys Leu Ile Trp Pro Asn Asn Tyr Gly Val Trp Val
140 145 150

gat gaa ttt gag gcc atg gat ttg ctt gat tgc ctt gat act act tgg 592
Asp Glu Phe Glu Ala Met Asp Leu Leu Asp Cys Leu Asp Thr Thr Trp
155 160 165

tct ggt gct gtt gtg cac att gat gat aat aca aag aag gat ctt aat 640
Ser Gly Ala Val Val His Ile Asp Asp Asn Thr Lys Lys Asp Leu Asn
170 175 180

```
aga cct tat ggg aga gtt aat agg aag ttg ctg aag tcg aaa atg ctg   688
Arg Pro Tyr Gly Arg Val Asn Arg Lys Leu Leu Lys Ser Lys Met Leu
185                 190                 195                 200

caa aaa tgc ata acc aat ggt gtt aag ttt cac caa gct aaa gtt att   736
Gln Lys Cys Ile Thr Asn Gly Val Lys Phe His Gln Ala Lys Val Ile
                205                 210                 215

aag gtt att cat gaa gag tcc aaa tct ttg ttg att tgc aat gat ggt   784
Lys Val Ile His Glu Glu Ser Lys Ser Leu Leu Ile Cys Asn Asp Gly
            220                 225                 230

gtg aca att cag gca gcc gtg gtt ctt gat gct acg ggg ttc tct agg   832
Val Thr Ile Gln Ala Ala Val Val Leu Asp Ala Thr Gly Phe Ser Arg
        235                 240                 245

tgt ctt gtc cag tat gat aag ccc tat aat cca ggt tac caa gtg gca   880
Cys Leu Val Gln Tyr Asp Lys Pro Tyr Asn Pro Gly Tyr Gln Val Ala
    250                 255                 260

tat gga ata cta gct gag gta gaa cag cac ccg ttt gat tta gac aag   928
Tyr Gly Ile Leu Ala Glu Val Glu Gln His Pro Phe Asp Leu Asp Lys
265                 270                 275                 280

atg gtt ttc atg gat tgg aga gat tcg cat ctg aac aac aat tcg cag   976
Met Val Phe Met Asp Trp Arg Asp Ser His Leu Asn Asn Asn Ser Gln
                285                 290                 295

ctc aaa gag gca aat agc aaa att cct act ttt ctt tat gcc atg ccc  1024
Leu Lys Glu Ala Asn Ser Lys Ile Pro Thr Phe Leu Tyr Ala Met Pro
            300                 305                 310

ttt tcg tca aac agg ata ttt ctt gaa gag act tcg cta gtg gcg cgg  1072
Phe Ser Ser Asn Arg Ile Phe Leu Glu Glu Thr Ser Leu Val Ala Arg
        315                 320                 325

cct gga gtg cca atg aaa gat atc cag gaa aga atg gtg gct aga tta  1120
Pro Gly Val Pro Met Lys Asp Ile Gln Glu Arg Met Val Ala Arg Leu
    330                 335                 340

aag cac tta ggc ata aaa gtt aaa agc att gaa gag gat gag cat tgt  1168
Lys His Leu Gly Ile Lys Val Lys Ser Ile Glu Glu Asp Glu His Cys
345                 350                 355                 360

gtc att ccg atg ggt ggg ccc ctt cca gtg ctt cct caa aga gtt gtt  1216
Val Ile Pro Met Gly Gly Pro Leu Pro Val Leu Pro Gln Arg Val Val
                365                 370                 375

gga ata ggt ggt acc gct ggg atg gtg cac cct tca act ggc tat atg  1264
Gly Ile Gly Gly Thr Ala Gly Met Val His Pro Ser Thr Gly Tyr Met
            380                 385                 390

gtg gca agg act tta gct gcg gct cct att gtt gca aat gca atc gtt  1312
Val Ala Arg Thr Leu Ala Ala Ala Pro Ile Val Ala Asn Ala Ile Val
        395                 400                 405

cga agc ctc agt tct gac aga agc att tca gga cac aaa ttg tct gct  1360
Arg Ser Leu Ser Ser Asp Arg Ser Ile Ser Gly His Lys Leu Ser Ala
    410                 415                 420

gaa gtt tgg aaa gat ttg tgg ccc ata gaa agg aga agg caa agg gag  1408
Glu Val Trp Lys Asp Leu Trp Pro Ile Glu Arg Arg Arg Gln Arg Glu
425                 430                 435                 440
```

```
ttc ttc tgt ttt ggt atg gat atc ctg ctc aaa ctt gac tta cct gcc   1456
Phe Phe Cys Phe Gly Met Asp Ile Leu Leu Lys Leu Asp Leu Pro Ala
                445                 450                 455

act agg agg ttt ttc gat gct ttt ttt gat ctg gag cct cgt tat tgg   1504
Thr Arg Arg Phe Phe Asp Ala Phe Phe Asp Leu Glu Pro Arg Tyr Trp
            460                 465                 470

cat ggt ttc tta tca tcg aga ttg ttt ctc ccc gag ctt tta gtt ttt   1552
His Gly Phe Leu Ser Ser Arg Leu Phe Leu Pro Glu Leu Leu Val Phe
        475                 480                 485

ggg ctt tct cta ttc tca cat gcc tct aat act tct agg cta gag atc   1600
Gly Leu Ser Leu Phe Ser His Ala Ser Asn Thr Ser Arg Leu Glu Ile
    490                 495                 500

atg gca aag gga act ctt cct ttg gtt aac atg atc aac aac ttg gta   1648
Met Ala Lys Gly Thr Leu Pro Leu Val Asn Met Ile Asn Asn Leu Val
505                 510                 515                 520

caa gat aca gat taaggtgacc atgatagtta taatgtgctt aataactcat      1700
Gln Asp Thr Asp

gcactaatcg tttataaaac acttcaaatt agttttgatg tttatagctt attacatgaa 1760

ccaaagctta tgatagacgt gctttggtat ttaagagttt cagccaaaaa aaaaaaaaaa 1820

aaaaaaaaaa                                                      1830
```

<210> 34
<211> 524
<212> PRT
<213> Citrus X paradisi

<400> 34

```
Met Leu Pro Phe Leu Ser Ser Leu Leu Asn Gly Val Thr Asp Asn Pro
  1               5                  10                  15

Cys Arg Lys Ala Met Asp Thr Leu Leu Lys Thr His Asn Lys Leu Glu
            20                  25                  30

Phe Leu Pro Gln Val His Gly Ala Leu Glu Lys Ser Ser Ser Leu Ser
        35                  40                  45

Ser Leu Lys Ile Gln Asn Gln Glu Leu Arg Phe Gly Leu Lys Lys Ser
    50                  55                  60

Arg Gln Lys Arg Asn Arg Ser Cys Phe Ile Lys Ala Ser Ser Ser Ala
65                  70                  75                  80

Leu Leu Glu Leu Val Pro Glu Thr Lys Lys Glu Asn Leu Glu Phe Glu
                85                  90                  95

Leu Pro Met Tyr Asp Pro Ser Lys Gly Leu Val Val Asp Leu Ala Val
            100                 105                 110

Val Gly Gly Gly Pro Ala Gly Leu Ala Val Ala Gln Gln Val Ser Gly
        115                 120                 125

Ala Gly Leu Ser Val Cys Ser Ile Asp Pro Ser Pro Lys Leu Ile Trp
    130                 135                 140

Pro Asn Asn Tyr Gly Val Trp Val Asp Glu Phe Glu Ala Met Asp Leu
145                 150                 155                 160

Leu Asp Cys Leu Asp Thr Thr Trp Ser Gly Ala Val Val His Ile Asp
                165                 170                 175

Asp Asn Thr Lys Lys Asp Leu Asn Arg Pro Tyr Gly Arg Val Asn Arg
            180                 185                 190
```

Lys Leu Leu Lys Ser Lys Met Leu Gln Lys Cys Ile Thr Asn Gly Val
195 200 205

Lys Phe His Gln Ala Lys Val Ile Lys Val Ile His Glu Glu Ser Lys
210 215 220

Ser Leu Leu Ile Cys Asn Asp Gly Val Thr Ile Gln Ala Ala Val Val
225 230 235 240

Leu Asp Ala Thr Gly Phe Ser Arg Cys Leu Val Gln Tyr Asp Lys Pro
245 250 255

Tyr Asn Pro Gly Tyr Gln Val Ala Tyr Gly Ile Leu Ala Glu Val Glu
260 265 270

Gln His Pro Phe Asp Leu Asp Lys Met Val Phe Met Asp Trp Arg Asp
275 280 285

Ser His Leu Asn Asn Asn Ser Gln Leu Lys Glu Ala Asn Ser Lys Ile
290 295 300

Pro Thr Phe Leu Tyr Ala Met Pro Phe Ser Ser Asn Arg Ile Phe Leu
305 310 315 320

Glu Glu Thr Ser Leu Val Ala Arg Pro Gly Val Pro Met Lys Asp Ile
325 330 335

Gln Glu Arg Met Val Ala Arg Leu Lys His Leu Gly Ile Lys Val Lys
340 345 350

Ser Ile Glu Glu Asp Glu His Cys Val Ile Pro Met Gly Gly Pro Leu
355 360 365

Pro Val Leu Pro Gln Arg Val Val Gly Ile Gly Gly Thr Ala Gly Met
370 375 380

Val His Pro Ser Thr Gly Tyr Met Val Ala Arg Thr Leu Ala Ala Ala
385 390 395 400

Pro Ile Val Ala Asn Ala Ile Val Arg Ser Leu Ser Ser Asp Arg Ser
405 410 415

Ile Ser Gly His Lys Leu Ser Ala Glu Val Trp Lys Asp Leu Trp Pro
420 425 430

Ile Glu Arg Arg Arg Gln Arg Glu Phe Phe Cys Phe Gly Met Asp Ile
435 440 445

Leu Leu Lys Leu Asp Leu Pro Ala Thr Arg Arg Phe Phe Asp Ala Phe
450 455 460

Phe Asp Leu Glu Pro Arg Tyr Trp His Gly Phe Leu Ser Ser Arg Leu
465 470 475 480

Phe Leu Pro Glu Leu Leu Val Phe Gly Leu Ser Leu Phe Ser His Ala
485 490 495

Ser Asn Thr Ser Arg Leu Glu Ile Met Ala Lys Gly Thr Leu Pro Leu
500 505 510

Val Asn Met Ile Asn Asn Leu Val Gln Asp Thr Asp
515 520

<210> 35
<211> 787
<212> DNA
<213> Citrus sinensis

<220>
<221> CDS
<222> (2)..(787)
<223> coding for epsilon-cyclase (partial)

<400> 35

```
t ctt gct ttg gct gca gaa tca gcg aag ttg gga tta aat gtt gga ctt  49
  Leu Ala Leu Ala Ala Glu Ser Ala Lys Leu Gly Leu Asn Val Gly Leu
    1               5                  10                  15

att ggc ccg gat ctc cct ttc aca aac aat tat ggt gtg tgg gaa gat   97
Ile Gly Pro Asp Leu Pro Phe Thr Asn Asn Tyr Gly Val Trp Glu Asp
            20                  25                  30

gaa ttt aga gat ctt gga ctt gaa ggg tgt atc gaa caa gtc tgg aga  145
Glu Phe Arg Asp Leu Gly Leu Glu Gly Cys Ile Glu Gln Val Trp Arg
        35                  40                  45

gac aca gtt gta tat att gat gaa gat gaa ccc atc ttg att ggt cgt  193
Asp Thr Val Val Tyr Ile Asp Glu Asp Glu Pro Ile Leu Ile Gly Arg
    50                  55                  60

gct tat gga cga gtt agt cga cat ttg ctt cat gaa gaa tta tta aga  241
Ala Tyr Gly Arg Val Ser Arg His Leu Leu His Glu Glu Leu Leu Arg
65                  70                  75                  80

agg tgt gtc gag tca ggt gta tca tat ctt agc tca aaa gtg gaa agc  289
Arg Cys Val Glu Ser Gly Val Ser Tyr Leu Ser Ser Lys Val Glu Ser
                85                  90                  95

att acg gaa tct acc agt ggt cat cgt ctt gta gct tgt gaa cat gat  337
Ile Thr Glu Ser Thr Ser Gly His Arg Leu Val Ala Cys Glu His Asp
            100                 105                 110

atg att gtc ccc tgc agg ctt gct act gtt gct tct gga gca gca tca  385
Met Ile Val Pro Cys Arg Leu Ala Thr Val Ala Ser Gly Ala Ala Ser
        115                 120                 125

ggg aag cta ttg gaa tat ggg gtg ggg ggt ccc aaa gtt tct gtc caa  433
Gly Lys Leu Leu Glu Tyr Gly Val Gly Gly Pro Lys Val Ser Val Gln
    130                 135                 140

aca gct tat ggt gtg gag gtt gag gtg gaa aat aat cca tat gat cca  481
Thr Ala Tyr Gly Val Glu Val Glu Val Glu Asn Asn Pro Tyr Asp Pro
145                 150                 155                 160

agc ctt atg gtt ttc atg gac tac aga gac tgt act aag caa gaa gtt  529
Ser Leu Met Val Phe Met Asp Tyr Arg Asp Cys Thr Lys Gln Glu Val
                165                 170                 175

cca tct ttt gaa tct gac aat cca aca ttt ctt tat gtc atg ccc atg  577
Pro Ser Phe Glu Ser Asp Asn Pro Thr Phe Leu Tyr Val Met Pro Met
            180                 185                 190

tct tca aca aga gtt ttc ttt gag gaa act tgt ttg gca tcg aaa gat  625
Ser Ser Thr Arg Val Phe Phe Glu Glu Thr Cys Leu Ala Ser Lys Asp
        195                 200                 205

ggc tta cgt ttt gac ata ttg aag aaa aag ctc atg gca agg tta gag  673
Gly Leu Arg Phe Asp Ile Leu Lys Lys Lys Leu Met Ala Arg Leu Glu
    210                 215                 220

aga ttg gga atc cag gtt ttg aaa act tat gaa gag gaa tgg tca tat  721
Arg Leu Gly Ile Gln Val Leu Lys Thr Tyr Glu Glu Glu Trp Ser Tyr
225                 230                 235                 240
```

att cca gtt ggt ggt tcc tta cca aat aca gaa caa aga aac ctc gca 769
Ile Pro Val Gly Gly Ser Leu Pro Asn Thr Glu Gln Arg Asn Leu Ala
245 250 255

tat ggt gct gct gct agc 787
Tyr Gly Ala Ala Ala Ser
260

<210> 36
<211> 262
<212> PRT
<213> Citrus sinensis

<400> 36
Leu Ala Leu Ala Ala Glu Ser Ala Lys Leu Gly Leu Asn Val Gly Leu
1 5 10 15

Ile Gly Pro Asp Leu Pro Phe Thr Asn Asn Tyr Gly Val Trp Glu Asp
20 25 30

Glu Phe Arg Asp Leu Gly Leu Glu Gly Cys Ile Glu Gln Val Trp Arg
35 40 45

Asp Thr Val Val Tyr Ile Asp Glu Asp Glu Pro Ile Leu Ile Gly Arg
50 55 60

Ala Tyr Gly Arg Val Ser Arg His Leu Leu His Glu Glu Leu Leu Arg
65 70 75 80

Arg Cys Val Glu Ser Gly Val Ser Tyr Leu Ser Ser Lys Val Glu Ser
85 90 95

Ile Thr Glu Ser Thr Ser Gly His Arg Leu Val Ala Cys Glu His Asp
100 105 110

Met Ile Val Pro Cys Arg Leu Ala Thr Val Ala Ser Gly Ala Ala Ser
115 120 125

Gly Lys Leu Leu Glu Tyr Gly Val Gly Gly Pro Lys Val Ser Val Gln
130 135 140

Thr Ala Tyr Gly Val Glu Val Glu Val Glu Asn Asn Pro Tyr Asp Pro
145 150 155 160

Ser Leu Met Val Phe Met Asp Tyr Arg Asp Cys Thr Lys Gln Glu Val
165 170 175

Pro Ser Phe Glu Ser Asp Asn Pro Thr Phe Leu Tyr Val Met Pro Met
180 185 190

Ser Ser Thr Arg Val Phe Phe Glu Glu Thr Cys Leu Ala Ser Lys Asp
195 200 205

Gly Leu Arg Phe Asp Ile Leu Lys Lys Lys Leu Met Ala Arg Leu Glu
210 215 220

Arg Leu Gly Ile Gln Val Leu Lys Thr Tyr Glu Glu Glu Trp Ser Tyr
225 230 235 240

Ile Pro Val Gly Gly Ser Leu Pro Asn Thr Glu Gln Arg Asn Leu Ala
245 250 255

Tyr Gly Ala Ala Ala Ser
260

<210> 37
<211> 2357

<212> DNA
<213> Spinacia oleracea

<220>
<221> CDS
<222> (264)..(1814)
<223> coding for episilon-cyclase

<400> 37
gcacgagaca ccacaaaacc attgaggaga gagaaagtca accaaatttc acaccttcca 60
cctccctctt ccatggccgc aacccaaacc cagccacctt caccgccgcc gtcgacagca 120
cactgaactt caccactaca aacttaaaaa aaatcttgga gaaatttgat tccgtaaaaa 180
tggagtatta ttgtctcgga gcttcgaaat tcgcaacaat ggcggtttct cctgcgctta 240
atcacgacaa tttgaggaat aaa atg gtt aaa caa cgc cag aat ttc cag acg 293
Met Val Lys Gln Arg Gln Asn Phe Gln Thr
1 5 10

ttt tgc ttt tgg agg ccg aat tct tcg aac gtt gta gta gaa tgt agt 341
Phe Cys Phe Trp Arg Pro Asn Ser Ser Asn Val Val Val Glu Cys Ser
15 20 25

agt cgt agg agt gga agt agt gtt ttg agg agt gcg aat agc gac agt 389
Ser Arg Arg Ser Gly Ser Ser Val Leu Arg Ser Ala Asn Ser Asp Ser
30 35 40

agt tgc gta att gcg cca gag gat ttt gcg aac gaa gaa gat ttc atc 437
Ser Cys Val Ile Ala Pro Glu Asp Phe Ala Asn Glu Glu Asp Phe Ile
45 50 55

aaa gct ggt ggt tcc gag ctt ctt tat gtt caa atg cag cag aat aaa 485
Lys Ala Gly Gly Ser Glu Leu Leu Tyr Val Gln Met Gln Gln Asn Lys
60 65 70

gct atg gat tgt tac tcc aaa att tcc gat aag ctg cgt caa ata tca 533
Ala Met Asp Cys Tyr Ser Lys Ile Ser Asp Lys Leu Arg Gln Ile Ser
75 80 85 90

gat gcc aat gaa ctg ctg gat atg gtg gtt att ggt tgt ggt cca gct 581
Asp Ala Asn Glu Leu Leu Asp Met Val Val Ile Gly Cys Gly Pro Ala
95 100 105

ggt cta gct ttg gct gca gaa tcg gct aaa ctt gga tta aaa gtt ggc 629
Gly Leu Ala Leu Ala Ala Glu Ser Ala Lys Leu Gly Leu Lys Val Gly
110 115 120

ctt gtt ggt cct gat ctt cct ttt acg aat aac tac ggc gtt tgg gaa 677
Leu Val Gly Pro Asp Leu Pro Phe Thr Asn Asn Tyr Gly Val Trp Glu
125 130 135

gat gaa ttc aga gca ttg gga ctt gga ggc tgt atc gag cac gtt tgg 725
Asp Glu Phe Arg Ala Leu Gly Leu Gly Gly Cys Ile Glu His Val Trp
140 145 150

cgt gat acc att gtg tat att gat gat gac aat cct ata tat att ggt 773
Arg Asp Thr Ile Val Tyr Ile Asp Asp Asp Asn Pro Ile Tyr Ile Gly
155 160 165 170

cga tct tat gga aaa gtc agc cgg caa tta ctt cac aag gaa ctg gtg 821
Arg Ser Tyr Gly Lys Val Ser Arg Gln Leu Leu His Lys Glu Leu Val
175 180 185

cac agg tgt ttg gag tca ggt gtc tct tat ctg aat gcg aaa gtg gaa 869
His Arg Cys Leu Glu Ser Gly Val Ser Tyr Leu Asn Ala Lys Val Glu
190 195 200

```
aat att atg gaa gga cct gat gga cat agg ctt gtt gct tgt gaa cgt   917
Asn Ile Met Glu Gly Pro Asp Gly His Arg Leu Val Ala Cys Glu Arg
        205                 210                 215

ggt gtc act att ccc tgc agg ctt gta act gtt gca tct gga gca gct   965
Gly Val Thr Ile Pro Cys Arg Leu Val Thr Val Ala Ser Gly Ala Ala
    220                 225                 230

tca ggg aaa ctt ctg gag tat gaa gtg ggt ggt cca agg gtt tgt gta  1013
Ser Gly Lys Leu Leu Glu Tyr Glu Val Gly Gly Pro Arg Val Cys Val
235                 240                 245                 250

caa aca gct tat ggt gtg gag gtg gag gtg gaa aac agt cct tat gat  1061
Gln Thr Ala Tyr Gly Val Glu Val Glu Val Glu Asn Ser Pro Tyr Asp
                255                 260                 265

ccc aat gtg atg gtg ttc atg gac tac aga gac tac act aaa ctg agc  1109
Pro Asn Val Met Val Phe Met Asp Tyr Arg Asp Tyr Thr Lys Leu Ser
            270                 275                 280

gtt caa tct ctg gag gca aag tat cca aca ttc ttg tat gca atg ccg  1157
Val Gln Ser Leu Glu Ala Lys Tyr Pro Thr Phe Leu Tyr Ala Met Pro
        285                 290                 295

ata tca cca act agg atc ttc ttt gag gag act tgc ttg gct tca gta  1205
Ile Ser Pro Thr Arg Ile Phe Phe Glu Glu Thr Cys Leu Ala Ser Val
    300                 305                 310

gat gca atg ccc ttt gac ctg ctc aag aaa aag ctt atg aca aga tta  1253
Asp Ala Met Pro Phe Asp Leu Leu Lys Lys Lys Leu Met Thr Arg Leu
315                 320                 325                 330

caa act atg ggt gtt cgt atc acc aaa ata tat gaa gag gag tgg tct  1301
Gln Thr Met Gly Val Arg Ile Thr Lys Ile Tyr Glu Glu Glu Trp Ser
                335                 340                 345

tat ata cct gtt ggt ggg tcc tta cca aat aca gag caa aga aac ctt  1349
Tyr Ile Pro Val Gly Gly Ser Leu Pro Asn Thr Glu Gln Arg Asn Leu
            350                 355                 360

gca ttt ggt gct gct gcg agc atg gtg cat cca gcc aca ggt tat tca  1397
Ala Phe Gly Ala Ala Ala Ser Met Val His Pro Ala Thr Gly Tyr Ser
        365                 370                 375

gtc gtg aga tca ctg tca gaa gct cca aag tat gct tct gca att gca  1445
Val Val Arg Ser Leu Ser Glu Ala Pro Lys Tyr Ala Ser Ala Ile Ala
    380                 385                 390

aac ttg atc aag aat gac ctg tca aaa aat gca ata ttg cgt cag agg  1493
Asn Leu Ile Lys Asn Asp Leu Ser Lys Asn Ala Ile Leu Arg Gln Arg
395                 400                 405                 410

agt gtg ggg aat atc tca atg caa gcc tgg aat act ctt tgg cca caa  1541
Ser Val Gly Asn Ile Ser Met Gln Ala Trp Asn Thr Leu Trp Pro Gln
                415                 420                 425

gaa agg aaa cgt cag aga gca ttc ttc ctg ttc gga cta tca ctt ata  1589
Glu Arg Lys Arg Gln Arg Ala Phe Phe Leu Phe Gly Leu Ser Leu Ile
            430                 435                 440

gtc cag ctt gat att gag ggt atc agg aca ttc ttc cgc acc ttc ttc  1637
Val Gln Leu Asp Ile Glu Gly Ile Arg Thr Phe Phe Arg Thr Phe Phe
        445                 450                 455
```

```
cga gtg cca aaa tgg atg tgg gag gga ttc ctc ggt tct aat ctc tct   1685
Arg Val Pro Lys Trp Met Trp Glu Gly Phe Leu Gly Ser Asn Leu Ser
    460                 465                 470

tca gct gat ctc ata ttg ttt gcc ttt tat atg ttc ttt att gct ccg   1733
Ser Ala Asp Leu Ile Leu Phe Ala Phe Tyr Met Phe Phe Ile Ala Pro
475                 480                 485                 490

aat gac ttg aga atg ggt ctt ata agg cat cta cta tct gat cct aca   1781
Asn Asp Leu Arg Met Gly Leu Ile Arg His Leu Leu Ser Asp Pro Thr
                495                 500                 505

ggg gcg acc atg ata aga acg tac ata aca cta taaaagtaat atgaaatgct 1834
Gly Ala Thr Met Ile Arg Thr Tyr Ile Thr Leu
            510                 515

cactcctttg tacatcatgc aaaattggta cgaattgact ggactatgca gtctaacttg 1894

gtgtaaaaaa aacacaatta ataaattttt tgtaggtgca gcctctatac ttgatattct 1954

cgattcagat ataatattgt cagtattctt cgttaaagat cagttgtttc tacaattcca 2014

gaggctcctg gaattggtgt tacccttcca tgtagctcat tgataaatgt tgagggtaga 2074

ggctttttct tagatgcttg cttgcagctt gctcatggat atattcagtt gttcagtaga 2134

cacgttaaca actactacag tgggggcatc attgatctgg accgggagag ctgagcatct 2194

atcacaggtt agccagctca actacgtagg tcaaccttga gccactccca aacatttttg 2254

cagctgatgg ggttcaccct gtaaggtgag tttcttacca actccaccaa cttatgttgg 2314

ttttaaattg ctactcgtct gttatgaagt agcaagctcg tgc                   2357
```

<210> 38
<211> 517
<212> PRT
<213> Spinacia oleracea

<400> 38

```
Met Val Lys Gln Arg Gln Asn Phe Gln Thr Phe Cys Phe Trp Arg Pro
  1               5                  10                  15

Asn Ser Ser Asn Val Val Val Glu Cys Ser Ser Arg Arg Ser Gly Ser
            20                  25                  30

Ser Val Leu Arg Ser Ala Asn Ser Asp Ser Ser Cys Val Ile Ala Pro
        35                  40                  45

Glu Asp Phe Ala Asn Glu Glu Asp Phe Ile Lys Ala Gly Gly Ser Glu
    50                  55                  60

Leu Leu Tyr Val Gln Met Gln Gln Asn Lys Ala Met Asp Cys Tyr Ser
65                  70                  75                  80

Lys Ile Ser Asp Lys Leu Arg Gln Ile Ser Asp Ala Asn Glu Leu Leu
                85                  90                  95

Asp Met Val Val Ile Gly Cys Gly Pro Ala Gly Leu Ala Leu Ala Ala
            100                 105                 110

Glu Ser Ala Lys Leu Gly Leu Lys Val Gly Leu Val Gly Pro Asp Leu
        115                 120                 125

Pro Phe Thr Asn Asn Tyr Gly Val Trp Glu Asp Glu Phe Arg Ala Leu
    130                 135                 140

Gly Leu Gly Gly Cys Ile Glu His Val Trp Arg Asp Thr Ile Val Tyr
145                 150                 155                 160

Ile Asp Asp Asp Asn Pro Ile Tyr Ile Gly Arg Ser Tyr Gly Lys Val
                165                 170                 175
```

Ser Arg Gln Leu Leu His Lys Glu Leu Val His Arg Cys Leu Glu Ser
180 185 190

Gly Val Ser Tyr Leu Asn Ala Lys Val Glu Asn Ile Met Glu Gly Pro
195 200 205

Asp Gly His Arg Leu Val Ala Cys Glu Arg Gly Val Thr Ile Pro Cys
210 215 220

Arg Leu Val Thr Val Ala Ser Gly Ala Ala Ser Gly Lys Leu Leu Glu
225 230 235 240

Tyr Glu Val Gly Gly Pro Arg Val Cys Val Gln Thr Ala Tyr Gly Val
245 250 255

Glu Val Glu Val Glu Asn Ser Pro Tyr Asp Pro Asn Val Met Val Phe
260 265 270

Met Asp Tyr Arg Asp Tyr Thr Lys Leu Ser Val Gln Ser Leu Glu Ala
275 280 285

Lys Tyr Pro Thr Phe Leu Tyr Ala Met Pro Ile Ser Pro Thr Arg Ile
290 295 300

Phe Phe Glu Glu Thr Cys Leu Ala Ser Val Asp Ala Met Pro Phe Asp
305 310 315 320

Leu Leu Lys Lys Lys Leu Met Thr Arg Leu Gln Thr Met Gly Val Arg
325 330 335

Ile Thr Lys Ile Tyr Glu Glu Glu Trp Ser Tyr Ile Pro Val Gly Gly
340 345 350

Ser Leu Pro Asn Thr Glu Gln Arg Asn Leu Ala Phe Gly Ala Ala Ala
355 360 365

Ser Met Val His Pro Ala Thr Gly Tyr Ser Val Val Arg Ser Leu Ser
370 375 380

Glu Ala Pro Lys Tyr Ala Ser Ala Ile Ala Asn Leu Ile Lys Asn Asp
385 390 395 400

Leu Ser Lys Asn Ala Ile Leu Arg Gln Arg Ser Val Gly Asn Ile Ser
405 410 415

Met Gln Ala Trp Asn Thr Leu Trp Pro Gln Glu Arg Lys Arg Gln Arg
420 425 430

Ala Phe Phe Leu Phe Gly Leu Ser Leu Ile Val Gln Leu Asp Ile Glu
435 440 445

Gly Ile Arg Thr Phe Phe Arg Thr Phe Phe Arg Val Pro Lys Trp Met
450 455 460

Trp Glu Gly Phe Leu Gly Ser Asn Leu Ser Ser Ala Asp Leu Ile Leu
465 470 475 480

Phe Ala Phe Tyr Met Phe Phe Ile Ala Pro Asn Asp Leu Arg Met Gly
485 490 495

Leu Ile Arg His Leu Leu Ser Asp Pro Thr Gly Ala Thr Met Ile Arg
500 505 510

Thr Tyr Ile Thr Leu
515

<210> 39
<211> 1378

<212> DNA
<213> Solanum tuberosum

<220>
<221> CDS
<222> (2)..(1147)
<223> coding for episilon-cyclase (partial)

<400> 39

```
t agc ggn nnn nag gat gag ttc aaa gat ctt ggt ctt caa gcc tgc att  49
  Ser Xaa Xaa Xaa Asp Glu Phe Lys Asp Leu Gly Leu Gln Ala Cys Ile
    1               5                  10                  15

gaa cat gtt tgg cgg gat acc att gta tat ctt gat gat gat gat cct   97
Glu His Val Trp Arg Asp Thr Ile Val Tyr Leu Asp Asp Asp Asp Pro
            20                  25                  30

att ctt att ggc cgt gcc tat gga aga gtt agt cgc cat tta ctg cac  145
Ile Leu Ile Gly Arg Ala Tyr Gly Arg Val Ser Arg His Leu Leu His
        35                  40                  45

gag gag tta ctc aaa agg tgt gtg gag gca ggt gtt ttg tat cta aac  193
Glu Glu Leu Leu Lys Arg Cys Val Glu Ala Gly Val Leu Tyr Leu Asn
    50                  55                  60

tcg aaa gtg gat agg att gtt gag gcc aca aat ggc cac agt ctt gta  241
Ser Lys Val Asp Arg Ile Val Glu Ala Thr Asn Gly His Ser Leu Val
65                  70                  75                  80

gag tgc gag ggt gat gtt gtg att ccc tgc agg ttt gtg act gtt gca  289
Glu Cys Glu Gly Asp Val Val Ile Pro Cys Arg Phe Val Thr Val Ala
                85                  90                  95

tcg gga gca gcc tcg ggg aaa ttc ttg cag tat gag ttg gga ggt cct  337
Ser Gly Ala Ala Ser Gly Lys Phe Leu Gln Tyr Glu Leu Gly Gly Pro
            100                 105                 110

aga gtt tct gtt caa aca gct tat gga gtg gaa gtt gag gtc gat aac  385
Arg Val Ser Val Gln Thr Ala Tyr Gly Val Glu Val Glu Val Asp Asn
        115                 120                 125

aat cca ttt gac ccg agc ctg atg gtt ttc atg gat tat aga gac tat  433
Asn Pro Phe Asp Pro Ser Leu Met Val Phe Met Asp Tyr Arg Asp Tyr
    130                 135                 140

gtc aga cac gac gct caa tct tta gaa gct aaa tat cca aca ttt ctc  481
Val Arg His Asp Ala Gln Ser Leu Glu Ala Lys Tyr Pro Thr Phe Leu
145                 150                 155                 160

tat gcc atg ccc atg tct cca aca cga gtc ttt ttc gag gaa act tgt  529
Tyr Ala Met Pro Met Ser Pro Thr Arg Val Phe Phe Glu Glu Thr Cys
                165                 170                 175

ttg gct tca aaa gat gca atg cca ttc gat ctg tta aag aaa aaa ttg  577
Leu Ala Ser Lys Asp Ala Met Pro Phe Asp Leu Leu Lys Lys Lys Leu
            180                 185                 190

atg tta cga ttg aac acc ctc ggt gta aga att aaa gaa att tat gag  625
Met Leu Arg Leu Asn Thr Leu Gly Val Arg Ile Lys Glu Ile Tyr Glu
        195                 200                 205

gag gaa tgg tct tac ata cca gtt gga gga tct ttg cca aat aca gaa  673
Glu Glu Trp Ser Tyr Ile Pro Val Gly Gly Ser Leu Pro Asn Thr Glu
    210                 215                 220
```

```
caa aaa aca ctt gca ttt ggt gct gct gct agc atg gtt cat cca gcc   721
Gln Lys Thr Leu Ala Phe Gly Ala Ala Ala Ser Met Val His Pro Ala
225                 230                 235                 240

aca ggt tat tca gtc gtc aga tca ctg tct gaa gct cca aaa tgc gcc   769
Thr Gly Tyr Ser Val Val Arg Ser Leu Ser Glu Ala Pro Lys Cys Ala
                245                 250                 255

ttc gtg ctt gca aat ata tta cga caa aat cat agc aag aat atg ctt   817
Phe Val Leu Ala Asn Ile Leu Arg Gln Asn His Ser Lys Asn Met Leu
            260                 265                 270

act agt tca agt acc ccg agt att tca act caa gct tgg aac act ctt   865
Thr Ser Ser Ser Thr Pro Ser Ile Ser Thr Gln Ala Trp Asn Thr Leu
        275                 280                 285

tgg cca caa gaa cga aaa cga caa aga tcg ttt ttc cta ttt gga ctg   913
Trp Pro Gln Glu Arg Lys Arg Gln Arg Ser Phe Phe Leu Phe Gly Leu
    290                 295                 300

gct ctg ata ttg cag ctg gat att gag ggg ata agg tca ttt ttc cgc   961
Ala Leu Ile Leu Gln Leu Asp Ile Glu Gly Ile Arg Ser Phe Phe Arg
305                 310                 315                 320

gcg ttc ttc cgt gtg cca aaa tgg atg tgg cag gga ttt ctt ggt tca   1009
Ala Phe Phe Arg Val Pro Lys Trp Met Trp Gln Gly Phe Leu Gly Ser
                325                 330                 335

agt ctt tct tna gca gac ctc atg tta ttt gcc ttc tac atg ttt att   1057
Ser Leu Ser Xaa Ala Asp Leu Met Leu Phe Ala Phe Tyr Met Phe Ile
            340                 345                 350

att gca cca aat gac atg aga aga ggc tta atc aga cat ctt tta tct   1105
Ile Ala Pro Asn Asp Met Arg Arg Gly Leu Ile Arg His Leu Leu Ser
        355                 360                 365

gat cct act ggt gca aca ttg ata aga act tat ctt aca ttt            1147
Asp Pro Thr Gly Ala Thr Leu Ile Arg Thr Tyr Leu Thr Phe
    370                 375                 380

tagagtaaat tcctcctaca atagttgttg aannagaggc ctcattactt cagattcata 1207

acagaaatcg cggtctctcg aggccttgta tataacattt tcactaggtt aatattgctt 1267

gaataagttg cacagtttca gtttttgtat ctgcttcttt tttgtccaag atcatgtatt 1327

ganccaattt atatacattg ccagtatata taaattttat aaaaaaaaaa a           1378
```

<210> 40
<211> 382
<212> PRT
<213> Solanum tuberosum

<400> 40

```
Ser Xaa Xaa Xaa Asp Glu Phe Lys Asp Leu Gly Leu Gln Ala Cys Ile
  1               5                  10                  15

Glu His Val Trp Arg Asp Thr Ile Val Tyr Leu Asp Asp Asp Asp Pro
             20                  25                  30

Ile Leu Ile Gly Arg Ala Tyr Gly Arg Val Ser Arg His Leu Leu His
         35                  40                  45

Glu Glu Leu Leu Lys Arg Cys Val Glu Ala Gly Val Leu Tyr Leu Asn
     50                  55                  60

Ser Lys Val Asp Arg Ile Val Glu Ala Thr Asn Gly His Ser Leu Val
 65                  70                  75                  80
```

```
Glu Cys Glu Gly Asp Val Val Ile Pro Cys Arg Phe Val Thr Val Ala
                85                  90                  95

Ser Gly Ala Ala Ser Gly Lys Phe Leu Gln Tyr Glu Leu Gly Gly Pro
            100                 105                 110

Arg Val Ser Val Gln Thr Ala Tyr Gly Val Glu Val Glu Val Asp Asn
        115                 120                 125

Asn Pro Phe Asp Pro Ser Leu Met Val Phe Met Asp Tyr Arg Asp Tyr
    130                 135                 140

Val Arg His Asp Ala Gln Ser Leu Glu Ala Lys Tyr Pro Thr Phe Leu
145                 150                 155                 160

Tyr Ala Met Pro Met Ser Pro Thr Arg Val Phe Phe Glu Glu Thr Cys
                165                 170                 175

Leu Ala Ser Lys Asp Ala Met Pro Phe Asp Leu Leu Lys Lys Lys Leu
            180                 185                 190

Met Leu Arg Leu Asn Thr Leu Gly Val Arg Ile Lys Glu Ile Tyr Glu
        195                 200                 205

Glu Glu Trp Ser Tyr Ile Pro Val Gly Gly Ser Leu Pro Asn Thr Glu
    210                 215                 220

Gln Lys Thr Leu Ala Phe Gly Ala Ala Ala Ser Met Val His Pro Ala
225                 230                 235                 240

Thr Gly Tyr Ser Val Val Arg Ser Leu Ser Glu Ala Pro Lys Cys Ala
                245                 250                 255

Phe Val Leu Ala Asn Ile Leu Arg Gln Asn His Ser Lys Asn Met Leu
            260                 265                 270

Thr Ser Ser Ser Thr Pro Ser Ile Ser Thr Gln Ala Trp Asn Thr Leu
        275                 280                 285

Trp Pro Gln Glu Arg Lys Arg Gln Arg Ser Phe Phe Leu Phe Gly Leu
    290                 295                 300

Ala Leu Ile Leu Gln Leu Asp Ile Glu Gly Ile Arg Ser Phe Phe Arg
305                 310                 315                 320

Ala Phe Phe Arg Val Pro Lys Trp Met Trp Gln Gly Phe Leu Gly Ser
                325                 330                 335

Ser Leu Ser Xaa Ala Asp Leu Met Leu Phe Ala Phe Tyr Met Phe Ile
            340                 345                 350

Ile Ala Pro Asn Asp Met Arg Arg Gly Leu Ile Arg His Leu Leu Ser
        355                 360                 365

Asp Pro Thr Gly Ala Thr Leu Ile Arg Thr Tyr Leu Thr Phe
    370                 375                 380
```

<210> 41
<211> 497
<212> DNA
<213> Daucus carota

<220>
<221> CDS
<222> (1)..(495)
<223> coding for episilon-cyclase (partial)

<400> 41

```
tat ggt gtt tgg gtg gat gaa ttt ata gat ctt gga ctt gaa ggg tgt   48
Tyr Gly Val Trp Val Asp Glu Phe Ile Asp Leu Gly Leu Glu Gly Cys
  1               5                  10                  15

att gag cat gtt tgg cgg gat act att gta tat ctt gat gat ggt gat   96
Ile Glu His Val Trp Arg Asp Thr Ile Val Tyr Leu Asp Asp Gly Asp
             20                  25                  30

cct att atg att ggc cgt gct tac gga aga gtt agt cgc cat ttg ctt  144
Pro Ile Met Ile Gly Arg Ala Tyr Gly Arg Val Ser Arg His Leu Leu
         35                  40                  45

cat gaa gaa ttg ctt aaa agg tgt gtc gag tca ggt gtt tcg tat ctt  192
His Glu Glu Leu Leu Lys Arg Cys Val Glu Ser Gly Val Ser Tyr Leu
     50                  55                  60

agc tca aaa gtt gaa aag att att gaa gct gga gat ggc cac agc ctg  240
Ser Ser Lys Val Glu Lys Ile Ile Glu Ala Gly Asp Gly His Ser Leu
 65                  70                  75                  80

gtt gag tgt gaa aat aat att gtc att cca tgc agg ctt gct act gtt  288
Val Glu Cys Glu Asn Asn Ile Val Ile Pro Cys Arg Leu Ala Thr Val
                 85                  90                  95

gca tct gga gca gct tct ggg aaa ctt ttg cag tat gag gtt ggg ggt  336
Ala Ser Gly Ala Ala Ser Gly Lys Leu Leu Gln Tyr Glu Val Gly Gly
            100                 105                 110

ccc aga gtt tct gtc caa aca gct tat ggt gtc gag gtt gag gtg gaa  384
Pro Arg Val Ser Val Gln Thr Ala Tyr Gly Val Glu Val Glu Val Glu
        115                 120                 125

aac aat cca tat gat ccc agt cta atg gtt ttc atg gat tac aga gat  432
Asn Asn Pro Tyr Asp Pro Ser Leu Met Val Phe Met Asp Tyr Arg Asp
    130                 135                 140

tat acc aaa caa aaa gtt cca ggc atg gag gca gaa tat cca act ttc  480
Tyr Thr Lys Gln Lys Val Pro Gly Met Glu Ala Glu Tyr Pro Thr Phe
145                 150                 155                 160

ctg tat gcc atg cca tt                                          497
Leu Tyr Ala Met Pro
                165
```

<210> 42
<211> 165
<212> PRT
<213> Daucus carota

<400> 42

```
Tyr Gly Val Trp Val Asp Glu Phe Ile Asp Leu Gly Leu Glu Gly Cys
  1               5                  10                  15

Ile Glu His Val Trp Arg Asp Thr Ile Val Tyr Leu Asp Asp Gly Asp
             20                  25                  30

Pro Ile Met Ile Gly Arg Ala Tyr Gly Arg Val Ser Arg His Leu Leu
         35                  40                  45

His Glu Glu Leu Leu Lys Arg Cys Val Glu Ser Gly Val Ser Tyr Leu
     50                  55                  60

Ser Ser Lys Val Glu Lys Ile Ile Glu Ala Gly Asp Gly His Ser Leu
 65                  70                  75                  80
```

```
Val Glu Cys Glu Asn Asn Ile Val Ile Pro Cys Arg Leu Ala Thr Val
                85                  90                  95

Ala Ser Gly Ala Ala Ser Gly Lys Leu Leu Gln Tyr Glu Val Gly Gly
            100                 105                 110

Pro Arg Val Ser Val Gln Thr Ala Tyr Gly Val Glu Val Glu Val Glu
        115                 120                 125

Asn Asn Pro Tyr Asp Pro Ser Leu Met Val Phe Met Asp Tyr Arg Asp
    130                 135                 140

Tyr Thr Lys Gln Lys Val Pro Gly Met Glu Ala Glu Tyr Pro Thr Phe
145                 150                 155                 160

Leu Tyr Ala Met Pro
                165
```

<210> 43
<211> 605
<212> DNA
<213> Daucus carota

<220>
<221> CDS
<222> (3)..(605)
<223> coding for episilon-cyclase (partial)

<400> 43

```
tc att ggc cgt gct tat gga aga tta gtc gcc att tgc ttc atg aag    47
   Ile Gly Arg Ala Tyr Gly Arg Leu Val Ala Ile Cys Phe Met Lys
     1               5                  10                  15

aat tgc tta aaa ggt gtg tcg agt cag gtg ttt cgt atc tta gct caa    95
Asn Cys Leu Lys Gly Val Ser Ser Gln Val Phe Arg Ile Leu Ala Gln
                20                  25                  30

aag ttg aaa aga tta ttg aag ctg gag atg gcc aca gcc tgg ttg agt   143
Lys Leu Lys Arg Leu Leu Lys Leu Glu Met Ala Thr Ala Trp Leu Ser
            35                  40                  45

gtg aaa ata ata ttg tca ttc cat gca ggc ttg cta ctg ttg cat ctg   191
Val Lys Ile Ile Leu Ser Phe His Ala Gly Leu Leu Leu Leu His Leu
        50                  55                  60

gag cag ctt ctg gga aac ttt tgc agt atg ggg ttg ggg gtc cca gag   239
Glu Gln Leu Leu Gly Asn Phe Cys Ser Met Gly Leu Gly Val Pro Glu
    65                  70                  75

ttt ctg tcc aaa cag ctt atg gtg tcg agg ttg agg tgg aaa cca atc   287
Phe Leu Ser Lys Gln Leu Met Val Ser Arg Leu Arg Trp Lys Pro Ile
80                  85                  90                  95

cca tat gat ccc agt cta atg gtt ttc atg gat tac aga gat tat acc   335
Pro Tyr Asp Pro Ser Leu Met Val Phe Met Asp Tyr Arg Asp Tyr Thr
                100                 105                 110

aaa caa aaa gtt cca ggc atg gag gca gaa tat cca aca ttt ctt tat   383
Lys Gln Lys Val Pro Gly Met Glu Ala Glu Tyr Pro Thr Phe Leu Tyr
            115                 120                 125

gtg atg ccc atg tcc cca aca agg att ttc ttt gag gag aca tgt ttg   431
Val Met Pro Met Ser Pro Thr Arg Ile Phe Phe Glu Glu Thr Cys Leu
        130                 135                 140
```

```
gct tca aaa gat gcg atg cca ttc gat cta ctg aag aaa aaa ctc atg   479
Ala Ser Lys Asp Ala Met Pro Phe Asp Leu Leu Lys Lys Lys Leu Met
    145                 150                 155

tca aga tta cag acg atg gga att cga gtt gcc aag aca tat gaa gag   527
Ser Arg Leu Gln Thr Met Gly Ile Arg Val Ala Lys Thr Tyr Glu Glu
160                 165                 170                 175

gaa tgg tct tat ata cct gtt ggg gga tct tta cct aat act gag caa   575
Glu Trp Ser Tyr Ile Pro Val Gly Gly Ser Leu Pro Asn Thr Glu Gln
                180                 185                 190

aag aat ctc gcc ttt ggt gct gcc gct aga                          605
Lys Asn Leu Ala Phe Gly Ala Ala Ala Arg
            195                 200
```

<210> 44
<211> 201
<212> PRT
<213> Daucus carota

<400> 44

```
Ile Gly Arg Ala Tyr Gly Arg Leu Val Ala Ile Cys Phe Met Lys Asn
  1               5                  10                  15

Cys Leu Lys Gly Val Ser Ser Gln Val Phe Arg Ile Leu Ala Gln Lys
            20                  25                  30

Leu Lys Arg Leu Leu Lys Leu Glu Met Ala Thr Ala Trp Leu Ser Val
        35                  40                  45

Lys Ile Ile Leu Ser Phe His Ala Gly Leu Leu Leu Leu His Leu Glu
    50                  55                  60

Gln Leu Leu Gly Asn Phe Cys Ser Met Gly Leu Gly Val Pro Glu Phe
65                  70                  75                  80

Leu Ser Lys Gln Leu Met Val Ser Arg Leu Arg Trp Lys Pro Ile Pro
                85                  90                  95

Tyr Asp Pro Ser Leu Met Val Phe Met Asp Tyr Arg Asp Tyr Thr Lys
            100                 105                 110

Gln Lys Val Pro Gly Met Glu Ala Glu Tyr Pro Thr Phe Leu Tyr Val
        115                 120                 125

Met Pro Met Ser Pro Thr Arg Ile Phe Phe Glu Glu Thr Cys Leu Ala
    130                 135                 140

Ser Lys Asp Ala Met Pro Phe Asp Leu Leu Lys Lys Lys Leu Met Ser
145                 150                 155                 160

Arg Leu Gln Thr Met Gly Ile Arg Val Ala Lys Thr Tyr Glu Glu Glu
                165                 170                 175

Trp Ser Tyr Ile Pro Val Gly Gly Ser Leu Pro Asn Thr Glu Gln Lys
            180                 185                 190

Asn Leu Ala Phe Gly Ala Ala Ala Arg
        195                 200
```

<210> 45
<211> 1697
<212> DNA
<213> Lycopersicon esculentum

<220>
<221> CDS
<222> (6)..(1583)
<223> coding for epsilon-cyclase

<400> 45

```
ttgaa atg gag tgt gtt gga gtt caa aat gtt gga gca atg gca gtt tta 50
      Met Glu Cys Val Gly Val Gln Asn Val Gly Ala Met Ala Val Leu
        1               5                  10                  15

acg cgt ccg aga ttg aac cgt tgg tcg gga gga gag tta tgc caa gaa   98
Thr Arg Pro Arg Leu Asn Arg Trp Ser Gly Gly Glu Leu Cys Gln Glu
                20                  25                  30

aaa agc atc ttt ttg gcg tat gag cag tat gaa agt aaa tgt aat agc   146
Lys Ser Ile Phe Leu Ala Tyr Glu Gln Tyr Glu Ser Lys Cys Asn Ser
            35                  40                  45

agt agt ggt agt gac agt tgt gta gtt gat aaa gaa gat ttt gct gat   194
Ser Ser Gly Ser Asp Ser Cys Val Val Asp Lys Glu Asp Phe Ala Asp
        50                  55                  60

gaa gaa gat tat ata aaa gcc ggt ggt tcg caa ctt gta ttt gtt caa   242
Glu Glu Asp Tyr Ile Lys Ala Gly Gly Ser Gln Leu Val Phe Val Gln
    65                  70                  75

atg cag cag aaa aaa gat atg gat cag cag tct aag ctt tct gat gag   290
Met Gln Gln Lys Lys Asp Met Asp Gln Gln Ser Lys Leu Ser Asp Glu
80                  85                  90                  95

tta cga caa ata tct gct gga caa acc gta ctg gat tta gtg gta atc   338
Leu Arg Gln Ile Ser Ala Gly Gln Thr Val Leu Asp Leu Val Val Ile
                100                 105                 110

ggc tgt ggt cct gct ggt ctt gct ctt gcc gcg gag tca gct aaa ttg   386
Gly Cys Gly Pro Ala Gly Leu Ala Leu Ala Ala Glu Ser Ala Lys Leu
            115                 120                 125

ggg ttg aac gtg ggg ctc gtt ggg cct gat ctt cct ttc aca aac aac   434
Gly Leu Asn Val Gly Leu Val Gly Pro Asp Leu Pro Phe Thr Asn Asn
        130                 135                 140

tat ggt gta tgg gag gac gag ttc aaa gat ctt ggt ctt caa gcc tgc   482
Tyr Gly Val Trp Glu Asp Glu Phe Lys Asp Leu Gly Leu Gln Ala Cys
    145                 150                 155

att gaa cat gtt tgg cgg gat acc att gta tat ctt gat gat gat gaa   530
Ile Glu His Val Trp Arg Asp Thr Ile Val Tyr Leu Asp Asp Asp Glu
160                 165                 170                 175

cct att ctt att ggc cgt gcc tat gga aga gtt agt cgc cat ttt ctg   578
Pro Ile Leu Ile Gly Arg Ala Tyr Gly Arg Val Ser Arg His Phe Leu
                180                 185                 190

cac gag gag tta ctc aaa agg tgt gtg gag gca ggt gtt ttg tat cta   626
His Glu Glu Leu Leu Lys Arg Cys Val Glu Ala Gly Val Leu Tyr Leu
            195                 200                 205

aac tcg aaa gtg gat agg att gtt gag gcc aca aat ggc cag agt ctt   674
Asn Ser Lys Val Asp Arg Ile Val Glu Ala Thr Asn Gly Gln Ser Leu
        210                 215                 220

gta gag tgc gaa ggt gat gtt gtg att ccc tgc agg ttt gtg act gtt   722
Val Glu Cys Glu Gly Asp Val Val Ile Pro Cys Arg Phe Val Thr Val
    225                 230                 235
```

```
gca tcg ggg gca gcc tcg ggg aaa ttc ttg cag tat gag ttg gga agt   770
Ala Ser Gly Ala Ala Ser Gly Lys Phe Leu Gln Tyr Glu Leu Gly Ser
240                 245                 250                 255

cct aga gtt tct gtt caa aca gct tat gga gtg gaa gtt gag gtt gat   818
Pro Arg Val Ser Val Gln Thr Ala Tyr Gly Val Glu Val Glu Val Asp
                260                 265                 270

aac aat cca ttt gac ccg agc ctg atg gtt ttc atg gat tat aga gat   866
Asn Asn Pro Phe Asp Pro Ser Leu Met Val Phe Met Asp Tyr Arg Asp
            275                 280                 285

tat ctc aga cac gac gct caa tct tta gaa gct aaa tat cca aca ttt   914
Tyr Leu Arg His Asp Ala Gln Ser Leu Glu Ala Lys Tyr Pro Thr Phe
        290                 295                 300

ctt tat gcc atg ccc atg tct cca aca cga gtc ttt ttc gag gaa act   962
Leu Tyr Ala Met Pro Met Ser Pro Thr Arg Val Phe Phe Glu Glu Thr
    305                 310                 315

tgt ttg gct tca aaa gat gca atg cca ttc gat ctg tta aag aaa aaa   1010
Cys Leu Ala Ser Lys Asp Ala Met Pro Phe Asp Leu Leu Lys Lys Lys
320                 325                 330                 335

ctg atg cta cga ttg aac acc ctt ggt gta aga att aaa gaa att tac   1058
Leu Met Leu Arg Leu Asn Thr Leu Gly Val Arg Ile Lys Glu Ile Tyr
                340                 345                 350

gag gag gaa tgg tct tac ata ccg gtt ggt gga tct ttg cca aat aca   1106
Glu Glu Glu Trp Ser Tyr Ile Pro Val Gly Gly Ser Leu Pro Asn Thr
            355                 360                 365

gaa caa aaa aca ctt gca ttt ggt gct gct gct agc atg gtt cat cca   1154
Glu Gln Lys Thr Leu Ala Phe Gly Ala Ala Ala Ser Met Val His Pro
        370                 375                 380

gcc aca ggt tat tca gtc gtc aga tca ctt tct gaa gct cca aaa tgc   1202
Ala Thr Gly Tyr Ser Val Val Arg Ser Leu Ser Glu Ala Pro Lys Cys
    385                 390                 395

gcc tct gta ctt gca aat ata tta cga caa cat tat agc aag aac atg   1250
Ala Ser Val Leu Ala Asn Ile Leu Arg Gln His Tyr Ser Lys Asn Met
400                 405                 410                 415

ctt acc agt tca agt atc ccg agt ata tca act caa gct tgg aac act   1298
Leu Thr Ser Ser Ser Ile Pro Ser Ile Ser Thr Gln Ala Trp Asn Thr
                420                 425                 430

ctt tgg cca caa gaa cga aaa cga caa aga tcg ttt ttc cta ttt gga   1346
Leu Trp Pro Gln Glu Arg Lys Arg Gln Arg Ser Phe Phe Leu Phe Gly
            435                 440                 445

ctg gct ctg ata ttg cag ctg gat att gag ggg ata agg tca ttt ttc   1394
Leu Ala Leu Ile Leu Gln Leu Asp Ile Glu Gly Ile Arg Ser Phe Phe
        450                 455                 460

cgc gca ttc ttc cgt gtg cca aaa tgg atg tgg cag gga ttt ctt ggt   1442
Arg Ala Phe Phe Arg Val Pro Lys Trp Met Trp Gln Gly Phe Leu Gly
    465                 470                 475

tca agt ctt tct tca gca gac ctc atg tta ttt gcc ttc tac atg ttt   1490
Ser Ser Leu Ser Ser Ala Asp Leu Met Leu Phe Ala Phe Tyr Met Phe
480                 485                 490                 495
```

att att gca cca aat gac atg aga aaa ggc ttg atc aga cat ctt tta 1538
Ile Ile Ala Pro Asn Asp Met Arg Lys Gly Leu Ile Arg His Leu Leu
500 505 510

tct gat cct act ggt gca aca ttg ata aga act tat ctt aca ttt 1583
Ser Asp Pro Thr Gly Ala Thr Leu Ile Arg Thr Tyr Leu Thr Phe
515 520 525

tagagtaaac tcctcctaca ataattgttg aatcagaggc ctcattactt cagattcata 1643

acagaaatcg cggtctctcg aggcattgta tataacattt tcactagctt aata 1697

<210> 46
<211> 526
<212> PRT
<213> Lycopersicon esculentum

<400> 46
Met Glu Cys Val Gly Val Gln Asn Val Gly Ala Met Ala Val Leu Thr
1 5 10 15

Arg Pro Arg Leu Asn Arg Trp Ser Gly Gly Glu Leu Cys Gln Glu Lys
20 25 30

Ser Ile Phe Leu Ala Tyr Glu Gln Tyr Glu Ser Lys Cys Asn Ser Ser
35 40 45

Ser Gly Ser Asp Ser Cys Val Val Asp Lys Glu Asp Phe Ala Asp Glu
50 55 60

Glu Asp Tyr Ile Lys Ala Gly Gly Ser Gln Leu Val Phe Val Gln Met
65 70 75 80

Gln Gln Lys Lys Asp Met Asp Gln Gln Ser Lys Leu Ser Asp Glu Leu
85 90 95

Arg Gln Ile Ser Ala Gly Gln Thr Val Leu Asp Leu Val Val Ile Gly
100 105 110

Cys Gly Pro Ala Gly Leu Ala Leu Ala Ala Glu Ser Ala Lys Leu Gly
115 120 125

Leu Asn Val Gly Leu Val Gly Pro Asp Leu Pro Phe Thr Asn Asn Tyr
130 135 140

Gly Val Trp Glu Asp Glu Phe Lys Asp Leu Gly Leu Gln Ala Cys Ile
145 150 155 160

Glu His Val Trp Arg Asp Thr Ile Val Tyr Leu Asp Asp Asp Glu Pro
165 170 175

Ile Leu Ile Gly Arg Ala Tyr Gly Arg Val Ser Arg His Phe Leu His
180 185 190

Glu Glu Leu Leu Lys Arg Cys Val Glu Ala Gly Val Leu Tyr Leu Asn
195 200 205

Ser Lys Val Asp Arg Ile Val Glu Ala Thr Asn Gly Gln Ser Leu Val
210 215 220

Glu Cys Glu Gly Asp Val Val Ile Pro Cys Arg Phe Val Thr Val Ala
225 230 235 240

Ser Gly Ala Ala Ser Gly Lys Phe Leu Gln Tyr Glu Leu Gly Ser Pro
245 250 255

Arg Val Ser Val Gln Thr Ala Tyr Gly Val Glu Val Glu Val Asp Asn
260 265 270

```
Asn Pro Phe Asp Pro Ser Leu Met Val Phe Met Asp Tyr Arg Asp Tyr
        275                 280                 285
Leu Arg His Asp Ala Gln Ser Leu Glu Ala Lys Tyr Pro Thr Phe Leu
    290                 295                 300
Tyr Ala Met Pro Met Ser Pro Thr Arg Val Phe Phe Glu Glu Thr Cys
305                 310                 315                 320
Leu Ala Ser Lys Asp Ala Met Pro Phe Asp Leu Leu Lys Lys Lys Leu
                325                 330                 335
Met Leu Arg Leu Asn Thr Leu Gly Val Arg Ile Lys Glu Ile Tyr Glu
            340                 345                 350
Glu Glu Trp Ser Tyr Ile Pro Val Gly Gly Ser Leu Pro Asn Thr Glu
        355                 360                 365
Gln Lys Thr Leu Ala Phe Gly Ala Ala Ala Ser Met Val His Pro Ala
    370                 375                 380
Thr Gly Tyr Ser Val Val Arg Ser Leu Ser Glu Ala Pro Lys Cys Ala
385                 390                 395                 400
Ser Val Leu Ala Asn Ile Leu Arg Gln His Tyr Ser Lys Asn Met Leu
                405                 410                 415
Thr Ser Ser Ser Ile Pro Ser Ile Ser Thr Gln Ala Trp Asn Thr Leu
            420                 425                 430
Trp Pro Gln Glu Arg Lys Arg Gln Arg Ser Phe Phe Leu Phe Gly Leu
        435                 440                 445
Ala Leu Ile Leu Gln Leu Asp Ile Glu Gly Ile Arg Ser Phe Phe Arg
    450                 455                 460
Ala Phe Phe Arg Val Pro Lys Trp Met Trp Gln Gly Phe Leu Gly Ser
465                 470                 475                 480
Ser Leu Ser Ser Ala Asp Leu Met Leu Phe Ala Phe Tyr Met Phe Ile
                485                 490                 495
Ile Ala Pro Asn Asp Met Arg Lys Gly Leu Ile Arg His Leu Leu Ser
            500                 505                 510
Asp Pro Thr Gly Ala Thr Leu Ile Arg Thr Tyr Leu Thr Phe
        515                 520                 525
```

<210> 47
<211> 510
<212> DNA
<213> Tagetes erecta

<220>
<221> misc_feature
<222> (1)..(510)
<223> coding for epsilon-cyclase specific probe

<400> 47

```
ggcacgaggc aaagcaaagg ttgtttgttg ttgttgttga gagacactcc aatccaaaca 60
gatacaaggc gtgactggat atttctctct cgttcctaac aacagcaacg aagaagaaaa 120
agaatcatta ctaacaatca atgagtatga gagctggaca catgacggca acaatggcgg 180
cttttacatg ccctaggttt atgactagca tcagatacac gaagcaaatt aagtgcaacg 240
ctgctaaaag ccagctagtc gttaaacaag agattgagga ggaagaagat tatgtgaaag 300
ccggtggatc ggagctgctt tttgttcaaa tgcaacagaa taagtccatg gatgcacagt 360
ctagcctatc ccaaaagctc ccaagggtac caataggagg aggaggagac agtaactgta 420
```

tactggattt ggttgtaatt ggttgtggtc ctgctggcct tgctcttgct ggagaatcag 480
ccaagctagg cttgaatgtc gcacttatcg 510

<210> 48
<211> 20
<212> DNA
<213> Artificial sequence

<220>
<223> Description of the artificial sequence:
oligonucleotide primer

<400> 48
ggcacgaggc aaagcaaagg 20

<210> 49
<211> 21
<212> DNA
<213> Artificial sequence

<220>
<223> Description of the artificial sequence:
oligonucleotide primer

<400> 49
cgataagtgc gacattcaag c 21

<210> 50
<211> 734
<212> DNA
<213> Tagetes erecta

<220>
<221> misc_feature
<222> (1)..(734)
<223> fragment of epsilon cyclase gene obtain by iPCR
comprising part of promoter region

<400> 50
ctaacaatca atgagtagag agctggacac atgacggcaa caatggcggc ttttacatgc 60
cctaggttta tgactagcat cagatacacg aagcaaatta agtgcaacgc tgctaaaagc 120
cagctagtcg ttaaacaaga gattgaggag gaagaagatt atgtgaaagc cggtggatcg 180
gagctgcttt ttgttcaaat gcaacagaat aagtccatgg atgcacagtc tagcctatcc 240
caaaaggtca ctccagactt aattgcttat aaataaataa atatgttttt taggaataat 300
gatatttaga tagattagct atcacctgtg ctgtggtgtg cagctcccaa gggtcttacc 360
gatagtaaaa tcgttagtta tgattaatac ttgggaggtg ggggattata ggctttgttg 420
tgagaatgtt gagaaagagg tttgacaaat cggtgtttga atgaggttaa atggagttta 480
attaaaataa agagaagaga aagattaaga gggtgatggg gatattaaag acggscaata 540
tagtgatgcc acgtagaaaa aggtaagtga aaacatacaa cgtggcttta aaagatggct 600
tggctgctaa tcaactcaac tcaactcata tcctatccat tcaaattcaa ttcaattcta 660
ttgaatgcaa agcaaagcaa aggttgtttg ttgttgttgt tgagagacac tccaatccaa 720
acagatacaa ggcg 734

<210> 51
<211> 280
<212> DNA
<213> Tagetes erecta

<220>
<221> misc_feature
<222> (1)..(280)
<223> fragment of epsilon cyclase gene obtain by
TAIL-PCR comprising part of promoter region

<400> 51
gtcgagtatg gagttcaatt aaaataaaga gaagaraaag attaagaggg tgatggggat 60
attaaagacg gccaatrtag tgatgccacg taagaaaaag gtaagtgaaa acatacaacg 120
tggctttaaa agatggcttg gctgctaatc aactcaactc aactcatatc ctatccattc 180
aaattcaatt caattctatt gaatgcaaag caaagcaaag caaaggttgt ttgttgttgt 240
tgttgagaga cactccaatc caaacagata caaggcgtga 280

<210> 52
<211> 23
<212> DNA
<213> Artificial sequence

<220>
<223> Description of the artificial sequence:
oligonucleotide primer

<400> 52
cgccttgtat ctgtttggat tgg 23

<210> 53
<211> 24
<212> DNA
<213> Artificial sequence

<220>
<223> Description of the artificial sequence:
oligonucleotide primer

<400> 53
ctaacaatca atgagtatga gagc 24

<210> 54
<211> 26
<212> DNA
<213> Artificial sequence

<220>
<223> Description of the artificial sequence:
oligonucleotide primer

<400> 54
agagcaaggc cagcaggacc acaacc 26

<210> 55
<211> 26
<212> DNA
<213> Artificial sequence

<220>
<223> Description of the artificial sequence:
oligonucleotide primer

<400> 55
ccttgggagc ttttgggata ggctag 26

<210> 56
<211> 26
<212> DNA
<213> Artificial sequence

<220>
<223> Description of the artificial sequence:
oligonucleotide primer

<400> 56
tcacgccttg tatctgtttg gattgg 26

<210> 57
<211> 15
<212> DNA
<213> Artifical sequence

<220>
<223> Description of the artificial sequence:
oligonucleotide primer

<400> 57
gtcgagtatg gagtt 15

<210> 58
<211> 734
<212> DNA
<213> Tagetes erecta

<220>
<221> misc_feature
<222> (1)..(734)
<223> coding for epsilon-cyclase genomic iPCR-fragment

<400> 58
ctaacaatca atgagtagag agctggacac atgacggcaa caatggcggc ttttacatgc 60
cctaggttta tgactagcat cagatacacg aagcaaatta agtgcaacgc tgctaaaagc 120
cagctagtcg ttaaacaaga gattgaggag gaagaagatt atgtgaaagc cggtggatcg 180
gagctgcttt ttgttcaaat gcaacagaat aagtccatgg atgcacagtc tagcctatcc 240
caaaaggtca ctccagactt aattgcttat aaataaataa atatgttttt taggaataat 300
gatatttaga tagattagct atcacctgtg ctgtggtgtg cagctcccaa gggtcttacc 360
gatagtaaaa tcgttagtta tgattaatac ttgggaggtg ggggattata ggctttgttg 420
tgagaatgtt gagaaagagg tttgacaaat cggtgtttga atgaggttaa atggagttta 480
attaaaataa agagaagaga aagattaaga gggtgatggg gatattaaag acggscaata 540
tagtgatgcc acgtagaaaa aggtaagtga aaacatacaa cgtggcttta aaagatggct 600
tggctgctaa tcaactcaac tcaactcata tcctatccat tcaaattcaa ttcaattcta 660
ttgaatgcaa agcaaagcaa aggttgtttg ttgttgttgt tgagagacac tccaatccaa 720
acagatacaa ggcg 734

<210> 59
<211> 25
<212> DNA
<213> Artificial sequence

<220>
<223> Description of the artificial sequence:
oligonucleotide primer

<400> 59
ctcgagagta aaatcgttag ttatg 25

<210> 60
<211> 29
<212> DNA
<213> Artificial sequence

<220>
<223> Description of the artificial sequence:
oligonucleotide primer

<400> 60
ccatggccat tgattgttag taatgattc 29

<210> 61
<211> 29

<212> DNA
<213> Artificial sequence

<220>
<223> Description of the artificial sequence:
oligonucleotide primer

<400> 61
ccatggtaat ttgcttcgtg tatctgatg 29

<210> 62
<211> 26
<212> DNA
<213> Artificial sequence

<220>
<223> Description of the artificial sequence:
oligonucleotide primer

<400> 62
ccatggcgct agcagcgaca gtaatg 26

<210> 63
<211> 23
<212> DNA
<213> Artificial sequence

<220>
<223> Description of the artificial sequence:
oligonucleotide primer

<400> 63
gatatccggt gtgagggaac tag 23

<210> 64
<211> 22
<212> DNA
<213> Artificial sequence

<220>
<223> Description of the artificial sequence:
oligonucleotide primer

<400> 64
gcaagctcga cagctacaaa cc 22

<210> 65
<211> 24
<212> DNA
<213> Artificial sequence

<220>
<223> Description of the artificial sequence:
oligonucleotide primer

<400> 65
gaagcatgca gctagcagcg acag 24

<210> 66
<211> 1795
<212> DNA
<213> Artificial sequence

<220>
<223> Description of the artificial sequence: coding for
ketolase - 35S terminator construct

<400> 66
ccatggcgct agcagcgaca gtaatgttgg agcagcttac cggaagcgct gaggcactca 60
aggagaagga gaaggaggtt gcaggcagct ctgacgtgtt gcgtacatgg gcgacccagt 120
actcgcttcc gtcagaggag tcagacgcgg cccgcccggg actgaagaat gcctacaagc 180
caccaccttc cgacacaaag ggcatcacaa tggcgctagc tgtcatcggc tcctgggccg 240
cagtgttcct ccacgccatt tttcaaatca agcttccgac ctccttggac cagctgcact 300
ggctgcccgt gtcagatgcc acagctcagc tggttagcgg cagcagcagc ctgctgcaca 360
tcgtcgtagt attctttgtc ctggagttcc tgtacacagg ccttttttatc accacgcatg 420
atgctatgca tggcaccatc gccatgagaa acaggcagct taatgacttc ttgggcagag 480
tatgcatctc cttgtacgcc tggtttgatt acaacatgct gcaccgcaag cattgggagc 540
accacaacca cactggcgag gtgggcaagg accctgactt ccacagggga aaccctggca 600
ttgtgccctg gtttgccagc ttcatgtcca gctacatgtc gatgtggcag tttgcgcgcc 660
tcgcatggtg gacggtggtc atgcagctgc tgggtgcgcc aatggcgaac ctgctggtgt 720
tcatggcggc cgcgcccatc ctgtccgcct tccgcttgtt ctactttggc acgtacatgc 780
cccacaagcc tgagcctggc gccgcgtcag gctcttcacc agccgtcatg aactggtgga 840
agtcgcgcac tagccaggcg tccgacctgg tcagctttct gacctgctac cacttcgacc 900
tgcactggga gcaccaccgc tggccctttg ccccctggtg ggagctgccc aactgccgcc 960
gcctgtctgg ccgaggtctg gttcctgcct agctggacac actgcagtgg gccctgctgc 1020
cagctgggca tgcctgcagg tcgacggatc cccgggaatt cggtacgctg aaatcaccag 1080
tctctctcta caaatctatc tctctctatt ttctccataa ataatgtgtg agtagtttcc 1140
cgataaggga aattagggtt cttatagggt ttcgctcatg tgttgagcat ataagaaacc 1200
cttagtatgt atttgtattt gtaaaatact tctatcaata aaatttctaa ttcctaaaac 1260
caaaatccag tactaaaatc cagatctcct aaagtcccta tagatctttg tcgtgaatat 1320
aaaccagaca cgagacgact aaacctggag cccagacgcc gttcgaagct agaagtaccg 1380
cttaggcagg aggccgttag ggaaaagatg ctaaggcagg gttggttacg ttgactcccc 1440
cgtaggtttg gtttaaatat gatgaagtgg acggaaggaa ggaggaagac aaggaaggat 1500
aaggttgcag gccctgtgca aggtaagaag atggaaattt gatagaggta cgctactata 1560
cttatactat acgctaaggg aatgcttgta tttataccct atacccccta ataacccctt 1620
atcaatttaa gaaataatcc gcataagccc ccgcttaaaa attggtatca gagccatgaa 1680
taggtctatg accaaaactc aagaggataa aacctcacca aaatacgaaa gagttcttaa 1740
ctctaaagat aaaagatctt tcaagatcaa aactagttcc ctcacaccgg atatc 1795

<210> 67
<211> 28
<212> DNA
<213> Artificial sequence

<220>
<223> Description of the artificial sequence:
oligonucleotide primer

<400> 67
gagctcactc actgatttcc attgcttg 28

<210> 68
<211> 37
<212> DNA
<213> Artificial sequence

<220>
<223> Description of the artificial sequence:
oligonucleotide primer

<400> 68
cgccgttaag tcgatgtccg ttgatttaaa cagtgtc 37

<210> 69
<211> 34
<212> DNA
<213> Artificial sequence

<220>
<223> Description of the artificial sequence:
oligonucleotide primer

<400> 69
atcaacggac atcgacttaa cggcgtttgt aaac 34

<210> 70
<211> 25
<212> DNA
<213> Artificial sequence

<220>
<223> Description of the artificial sequence:
oligonucleotide primer

<400> 70
taagcttttt gttgaagaga tttgg 25

<210> 71
<211> 28
<212> DNA
<213> Artificial sequence

<220>
<223> Description of the artificial sequence:
oligonucleotide primer

<400> 71
gtcgactacg taagtttctg cttctacc 28

<210> 72
<211> 26
<212> DNA
<213> Artificial sequence

<220>
<223> Description of the artificial sequence:
oligonucleotide primer

<400> 72
ggatccggtg atacctgcac atcaac 26

<210> 73
<211> 28
<212> DNA
<213> Artificial sequence

<220>
<223> Description of the artificial sequence:
oligonucleotide primer

<400> 73
aagcttaccg atagtaaaat cgttagtt 28

<210> 74
<211> 31
<212> DNA
<213> Artifical sequence

<220>
<223> Description of the artificial sequence:
oligonucleotide primer

<400> 74
ctcgagctta ccgatagtaa aatcgttagt t 31

<210> 75
<211> 28
<212> DNA
<213> Artificial sequence

<220>
<223> Description of the artificial sequence:
oligonucleotide primer

<400> 75
gtcgacaaca acaacaaaca acctttgc 28

<210> 76
<211> 28
<212> DNA
<213> Artificial sequence

<220>
<223> Description of the artificial sequence:
oligonucleotide primer

<400> 76
ggatccaaca acaacaaaca acctttgc 28

<210> 77
<211> 777
<212> DNA
<213> Arabidopsis thaliana

<220>
<221> promoter
<222> (1)..(777)
<223> modified version of the AP3 promoter

<400> 77
gagctcactc actgatttcc attgcttgaa aattgatgat gaactaagat caatccatgt 60
tagtttcaaa acaacagtaa ctgtggccaa cttagttttg aaacaacact aactggtcga 120
agcaaaaaga aaaaagagtt tcatcatata tctgatttga tggactgttt ggagttagga 180
ccaaacatta tctacaaaca aagacttttc tcctaacttg tgattccttc ttaaacccta 240
ggggtaatat tctatttttcc aaggatcttt agttaaaggc aaatccggga aattattgta 300
atcatttggg gaaacatata aaagatttga gttagatgga agtgacgatt aatccaaaca 360
tatatatctc tttcttctta tttcccaaat taacagacaa aagtagaata ttggctttta 420
acaccaatat aaaaacttgc ttcacaccta aacacttttg tttactttag ggtaagtgca 480
aaaagccaac caaatccacc tgcactgatt tgacgtttac aaacgccgtt aagtcgatgt 540
ccgttgattt aaacagtgtc ttgtaattaa aaaaatcagt ttacataaat ggaaaattta 600
tcacttagtt ttcatcaact tctgaactta cctttcatgg attaggcaat actttccatt 660
tttagtaact caagtggacc ctttacttct tcaactccat ctctctcttt ctatttcact 720
tctttcttct cattatatct cttgtcctct ccaccaaatc tcttcaacaa aaagctt 777

<210> 78
<211> 212
<212> DNA
<213> Solanum tuberosum

<220>
<221> intron
<222> (1)..(212)
<223> PIV2 intron of ST-LS1 gene

<400> 78
gtcgactacg taagtttctg cttctacctt tgatatatat ataataatta tcattaatta 60
gtagtaatat aatatttcaa atatttttttt caaaataaaa gaatgtagta tatagcaatt 120

gcttttctgt agtttataag tgtgtatatt ttaatttata acttttctaa tatatgacca 180
aaatttgttg atgtgcaggt atcaccggat cc 212

<210> 79
<211> 358
<212> DNA
<213> Tagetes erecta

<220>
<221> misc_feature
<222> (1)..(358)
<223> coding for sense-strand of epsilon cyclase
promoter directed dsRNA

<400> 79
aagcttaccg atagtaaaat cgttagttat gattaatact tgggaggtgg gggattatag 60
gctttgttgt gagaatgttg agaaagaggt ttgacaaatc ggtgtttgaa tgaggttaaa 120
tggagtttaa ttaaaataaa gagaagagaa agattaagag ggtgatgggg atattaaaga 180
cggccaatat agtgatgcca cgtagaaaaa ggtaagtgaa aacatacaac gtggctttaa 240
aagatggctt ggctgctaat caactcaact caactcatat cctatccatt caaattcaat 300
tcaattctat tgaatgcaaa gcaaagcaaa gcaaaggttg tttgttgttg ttgtcgac 358

<210> 80
<211> 361
<212> DNA
<213> Tagetes erecta

<220>
<221> misc_feature
<222> (1)..(361)
<223> coding for antisense-strand of epsilon cyclase
promoter directed dsRNA

<400> 80
ctcgagctta ccgatagtaa aatcgttagt tatgattaat acttgggagg tgggggatta 60
taggctttgt tgtgagaatg ttgagaaaga ggtttgacaa atcggtgttt gaatgaggtt 120
aaatggagtt taattaaaat aaagagaaga gaaagattaa gagggtgatg gggatattaa 180
agacggccaa tatagtgatg ccacgtagaa aaaggtaagt gaaaacatac aacgtggctt 240
taaaagatgg cttggctgct aatcaactca actcaactca tatcctatcc attcaaattc 300
aattcaattc tattgaatgc aaagcaaagc aaagcaaagg ttgtttgttg ttgttggatc 360
c 361

<210> 81
<211> 1537
<212> DNA
<213> Cucumis sativus

<220>
<221> promoter
<222> (1)..(1537)
<223> promoter of chromoplast-specific
carotenoid-associated protein (CHRC)

<400> 81
gagctctaca aattagggtt actttattca ttttcatcca ttctctttat tgttaaattt 60
tgtacattta ttcaataata ttatatgttt attacaaatt ctcactttct tattcatacc 120
tattcactca agcctttacc atcttccttt tctatttcaa tactatttct acttcatttt 180
tcacgttttt aacatctttc tttatttctt gtccacttcg tttagggatg cctaatgtcc 240
caaatttcat ctctcgtagt aacacaaaac caatgtaatg ctacttctct ctacattttt 300
aatacaaata aagtgaaaca aaatatctat aaataaacaa atatatatat tttgttagac 360
gctgtctcaa cccatcaatt aaaaaatttt gttatatttc tactttacct actaaatttg 420
tttctcatat ttacctttta acccccacaa aaaaaaatta taaaaaagaa agaaaaaagc 480

taaaccctat ttaaatagct aactataaga tcttaaaatt atcctcatca gtgtatagtt 540
taattggtta ttaacttata acattatata tctatgacat atactctctc ctagctattt 600
ctcacatttt ttaacttaag aaaatagtca taacatagtc taaaattcaa acatccacat 660
gctctaattt gattaacaaa aagttagaaa tatttattta aataaaaaag actaataaat 720
atataaaatg aatgttcata cgcagaccca tttagagatg agtatgcttt cacatgctga 780
gattattttc aaaactaagg ttgtagcaat attaaatcaa taaaattatt ataaataaca 840
aaattaacct gctcgtgttt gctgtatatg ggaggctaca aaataaatta aactaaagat 900
gattatgttt tagacatttt ttctatctgt attagtttat acatattaat tcaggagctg 960
cacaacccaa ttctattttc gttccttggt ggctgggttt ctcacaaggt tcaatagtca 1020
atattaggtt ttattggact tttaatagta tcaaacaaat ctatgtgtga acttaaaaat 1080
tgtattaaat atttagggta acctgttgcc gtttttagaa taatgtttct tcttaataca 1140
cgaaagcgta ttgtgtattc attcatttgg cgcctcacat gcttcggttg gctcgcttta 1200
gtctctgcct tctttgtata ttgtactccc cctcttccta tgccacgtgt tctgagctta 1260
acaagccacg ttgcgtgcca ttgccaaaca agtcatttta acttcacaag gtccgatttg 1320
acctccaaaa caacgacaag tttccgaaca gtcgcgaaga tcaagggtat aatcgtcttt 1380
ttgaattcta tttctcttta tttaatagtc cctctcgtgt gatagttttt aaaagatttt 1440
taaaacgtag ctgctgttta agtaaatccc agtccttcag tttgtgcttt tgtgtgtttt 1500
gtttctctga tttacggaat ttggaaataa taagctt 1537

<210> 82
<211> 22
<212> DNA
<213> Artificial sequence

<220>
<223> Description of the artificial sequence:
oligonucleotide primer

<400> 82
gagctctaca aattagggtt ac 22

<210> 83
<211> 23
<212> DNA
<213> Artificial sequence

<220>
<223> Description of the artificial sequence:
oligonucleotide primer

<400> 83
aagcttatta tttccaaatt ccg 23

We claim:

1. A method for the transgenic expression of nucleic acid sequences in the flower of plants, including the following steps

   I. introduction of a transgenic expression cassette into plant cells, where the transgenic expression cassette comprises at least the following elements

      a) at least one promoter sequence of a gene coding for an ε-cyclase, and

      b) at least one further nucleic acid sequence, and

      c) where appropriate further genetic control elements,

      where at least one of said promoter sequences and one further nucleic acid sequence are functionally linked together, and the further nucleic acid sequence is heterologous in relation to the promoter sequence or the plant cell, and

   II. selection of transgenic cells which comprise said expression cassette stably integrated into the genome, and

   III. regeneration of complete plants from said transgenic cells, where at least one of the further nucleic acid sequences is expressed in the flower.

2. The method as claimed in claim 1, where the promoter sequence of a gene coding for an ε-cyclase is a sequence selected from the group of sequences consisting of

   i) the promoter sequence of the ε-cyclase from Tagetes erecta as shown in SEQ ID NO: 1, the ε-cyclase from Arabidopsis thaliana as shown in SEQ ID NO: 7, the ε-cyclase from Oryza sativa as shown in SEQ ID NO: 8, and

   ii) functional equivalents of the promoter sequences as shown in SEQ ID NO: 1, 7 or 8 having substantially the same promoter activity as the promoter of the ε-cyclases as shown in SEQ ID NO: 1, 7 or 8 and

iii) functionally equivalent fragments of the sequences under i) or ii) having substantially the same promoter activity as the promoter of ε-cyclases as shown in SEQ ID NO: 1, 7 or 8.

3. A method for identifying and/or isolating promoters of genes which code for an ε-cyclase, where at least one nucleic acid sequence or a part thereof is employed in the identification and/or isolation, where said nucleic acid sequence codes for an amino acid sequence which comprises at least one sequence as shown in SEQ ID NO: 17, 18, 19, 20, 21 or 22 or a variation of these sequences.

4. The method as claimed in claim 3, where said nucleic acid sequence comprises a sequence as shown in SEQ ID NO: 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43 or 45.

5. The method as claimed in either of claims 3 or 4, where the method is carried out with use of the polymerase chain reaction, and said nucleic acid sequence or a part thereof is employed as primer.

6. A method for producing a transgenic expression cassette with specificity for the flower of plants, including the following steps:

I. isolation of a promoter sequence, where at least one nucleic acid sequence or a part thereof is employed in the isolation, where said nucleic acid sequence codes for an amino acid sequence which comprises at least one sequence as shown in SEQ ID NO: 17, 18, 19, 20, 21 or 22 or a variation of these sequence motifs.

II. functional linkage of said promoter sequence to a further nucleic acid sequence, where said nucleic acid sequence is heterologous in relation to the promoter.

7. The method as claimed in claim 6, where said nucleic acid sequence includes a sequence as shown in SEQ ID NO: 23, 25, 27, 29, 29, 31, 33, 35, 37, 39, 41, 43 or 45.

8. The method as claimed in either of claims 6 or 7, where the method is carried out with use of the polymerase chain reaction, and said nucleic acid sequence or a part thereof is employed as primer.

9. A transgenic expression cassette for the targeted transgenic expression of nucleic acid sequences in the flower of plants, including

a) at least one promoter sequence of a gene coding for an ε-cyclase, and

b) at least one further nucleic acid sequence, and

c) where appropriate further genetic control elements,

where at least one promoter sequence and one further nucleic acid sequence are functionally linked together, and the further nucleic acid sequence is heterologous in relation to the promoter sequence.

10. The transgenic expression cassette as claimed in claim 9, where the promoter sequence of a gene coding for an ε-cyclase is a sequence selected from the group of sequences consisting of

i) the promoter sequence of the ε-cyclase from Tagetes erecta as shown in SEQ ID NO: 1, the ε-cyclase from Arabidopsis thaliana as shown in SEQ ID NO: 7, the ε-cyclase from Oryza sativa as shown in SEQ ID NO: 8, and

ii) functional equivalents of the promoter sequences as shown in SEQ ID NO: 1, 7 or 8 having substantially the same promoter activity as the promoter of the ε-cyclases as shown in SEQ ID NO: 1, 7 or 8 and

iii) functionally equivalent fragments of the sequences under i) or ii) having substantially the same promoter activity as the promoter of ε-cyclases as shown in SEQ ID NO: 1, 7 or 8.

11. The transgenic expression cassette as claimed in claim 9 or 10, where the nucleic acid sequence to be expressed transgenically enables

a) the expression of a protein encoded by said nucleic acid sequence, or

b) the expression of a sense RNA, antisense RNA or double-stranded RNA encoded by said nucleic acid sequence.

12. An isolated nucleic acid sequence comprising

a) the Tagetes erecta ε-cyclase promoter as shown in SEQ ID NO: 1 or

b) a functionally equivalent fragment of a) with substantially the same promoter activity as a).

13. The isolated nucleic acid sequence as claimed in claim 12, including, in the 3' orientation to the Tagetes erecta ε-cyclase promoter as shown in SEQ ID NO: 1 or a functionally equivalent fragment of the aforementioned, a sequence coding for a 5'-untranslated region and/or a transit peptide.

14. The isolated nucleic acid sequence as claimed in claim 12 or 13 including a sequence described by SEQ ID NO: 2 or 3.

15. A double-stranded RNA molecule comprising

a) a sense RNA strand comprising at least one ribonucleotide sequence which is substantially identical to at least part of a nucleic acid sequence coding for the promoter region of an ε-cyclase, and

b) an antisense RNA strand which is substantially complementary to the RNA sense strand under a).

16. The double-stranded RNA molecule as claimed in claim 15, where the promoter region of the ε-cyclase comprises a sequence selected from the sequences as shown in SEQ ID NO: 1, 7 or 8.

17. A ribonucleic acid molecule comprising

a) at least one ribonucleotide sequence which is substantially identical to at least one part of a nucleic acid sequence coding for the promoter region of an ε-cyclase, and

b) at least one further ribonucleotide sequence which is substantially complementary to at least one part of the ribonucleotide sequence under a),

where a) and b) are connected together covalently, and further functional elements may be located where appropriate between a) and b).

18. The ribonucleic acid molecule as claimed in claim 17, where the promoter region of the ε-cyclase includes a sequence selected from the sequences as shown in SEQ ID NO: 1, 7 or 8.

19. A transgenic expression cassette, comprising

a) at least one promoter functional in plants, and

b) at least one nucleic acid sequence coding for a double-stranded RNA molecule as claimed in either of claims 15 or 16 or coding for a ribonucleic acid molecule as claimed in either of claims 17 or 18,

where at least one of said promoters and at least one of said nucleic acid sequences are functionally linked together, and the promoter is heterologous in relation to the nucleic acid sequence.

20. The transgenic expression cassette as claimed in claim 19, where the promoter is a promoter having specificity for the flower of plants.

21. A transgenic expression vector comprising a nucleic acid sequence as claimed in any of claims 12 to 14 or a transgenic expression cassette as claimed in any of claims 9, 10, 11, 19 or 20.

22. A transgenic organism comprising a nucleic acid sequence as claimed in any of claims 12 to 14, a double-stranded RNA as claimed in claim 15 or 16, a ribonucleotide sequence as claimed in claim 17 or 18, a transgenic expression cassette as claimed in any of claims 9, 10, 11, 19 or 20 or a transgenic expression vector as claimed in claim 21.

23. The transgenic organism as claimed in claim 22 selected from the group consisting of bacteria, yeasts, fungi, animal and plant organisms.

24. The transgenic organism as claimed in claim 22 selected from the group consisting of bacteria, yeasts, fungi, non-human animal and plant organisms or cells, cell cultures, parts, tissues, organs or propagation material derived therefrom.

25. The transgenic organism as claimed in claim 23 or 24 selected from the group of agricultural crop plants.

26. The use of an isolated nucleic acid sequence as claimed in any of claims 12 to 14, of a double-stranded RNA as claimed in claim 15 or 16, of a ribonucleotide sequence as claimed in claim 17 or 18, of a transgenic expression cassette as claimed in any of claims 9, 10, 11, 19 or 20, of a transgenic expression vector as claimed in claim 21 of a transgenic organism as claimed in any of claims 23 to 25 or cell cultures, parts, organs, tissues or transgenic propagation material derived therefrom in methods for the transgenic expression of nucleic acids or proteins.

27. The use of an isolated nucleic acid sequence as claimed in any of claims 12 to 14, of a double-stranded RNA as claimed in claim 15 or 16, of a ribonucleotide sequence as claimed in claim 17 or 18, of a transgenic expression cassette as claimed in any of claims 9, 10, 11, 19 or 20, of a transgenic expression vector as claimed in claim 21, or of a transgenic organism as claimed in any of claims 23 to 25 or cell cultures, parts, organs, tissues or transgenic propagation material derived therefrom for producing human or animal foods, seeds, pharmaceuticals or fine chemicals.

28. A method for producing human or animal foods, seeds, pharmaceuticals or fine chemicals, where a transgenic organism as claimed in any of claims 23 to 25 is cultured, and the desired human or animal foods, seeds, pharmaceuticals or fine chemical is produced and/or isolated using said organism.

29. A method for producing ketocarotenoids, where the mRNA amount and/or activity of at least one ε-cyclase is reduced by introducing at least one double-stranded RNA as claimed in claim 15 or 16, one ribonucleotide sequence as claimed in claim 17 or 18 or one transgenic expression cassette as claimed in either of claims 19 or 20.

Application number/numéro de demande: EP03/08394

Figures: 1

Pages:

DRW-IP

Unscannable items
received with this application
(Request original documents in File Prep. Section on the 10th Floor)

Documents reçus avec cette demande ne pouvant être balayés
(Commander les documents originaux dans la section de préparation des dossiers au 10ième étage)

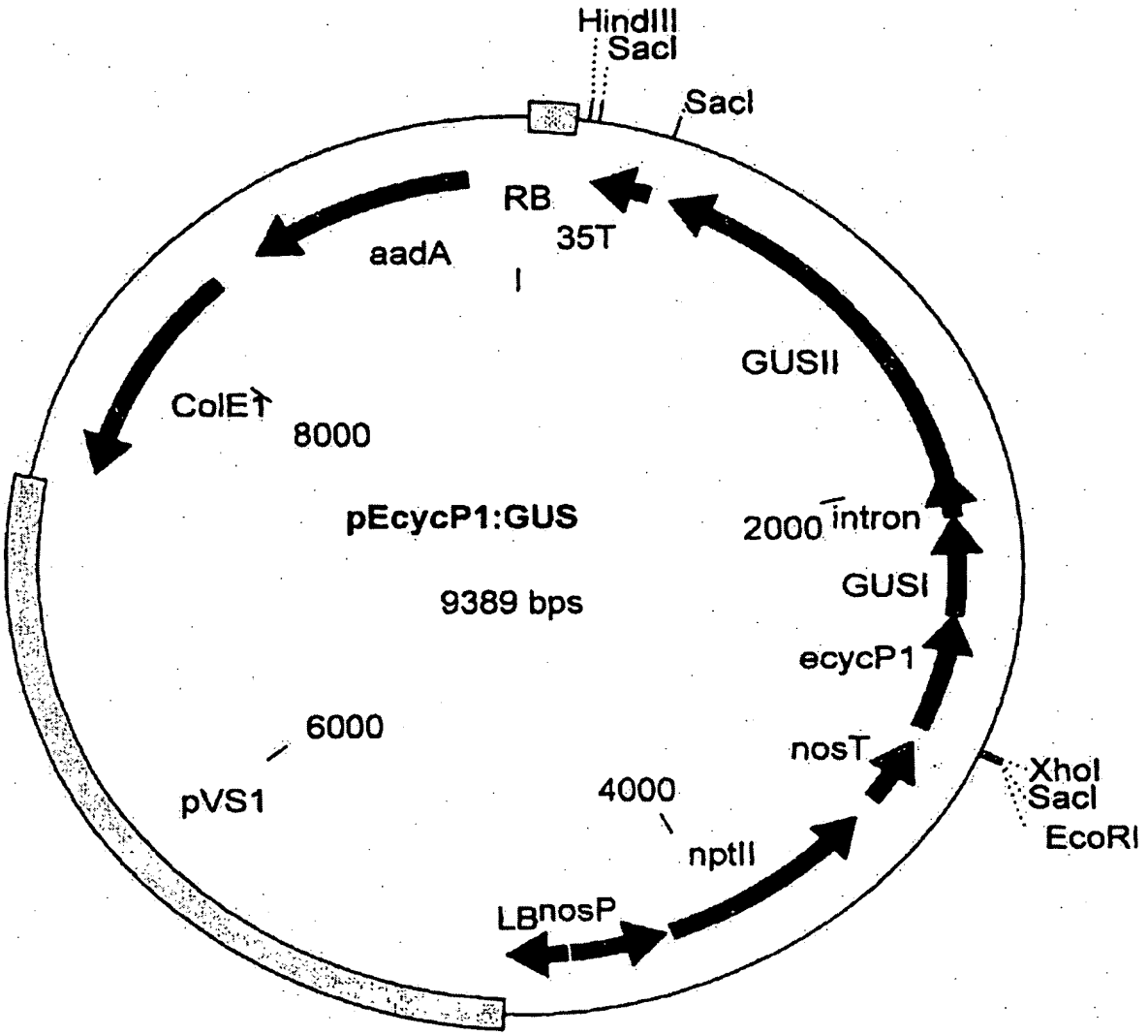

HindIII
SacI
SacI
RB
35T
aadA
GUSII
ColE1
8000
pEcycP1:GUS
2000
intron
9389 bps
GUSI
ecycP1
6000
nosT
XhoI
SacI
EcoRI
pVS1
4000
nptII
LB
nosP

Fig.2

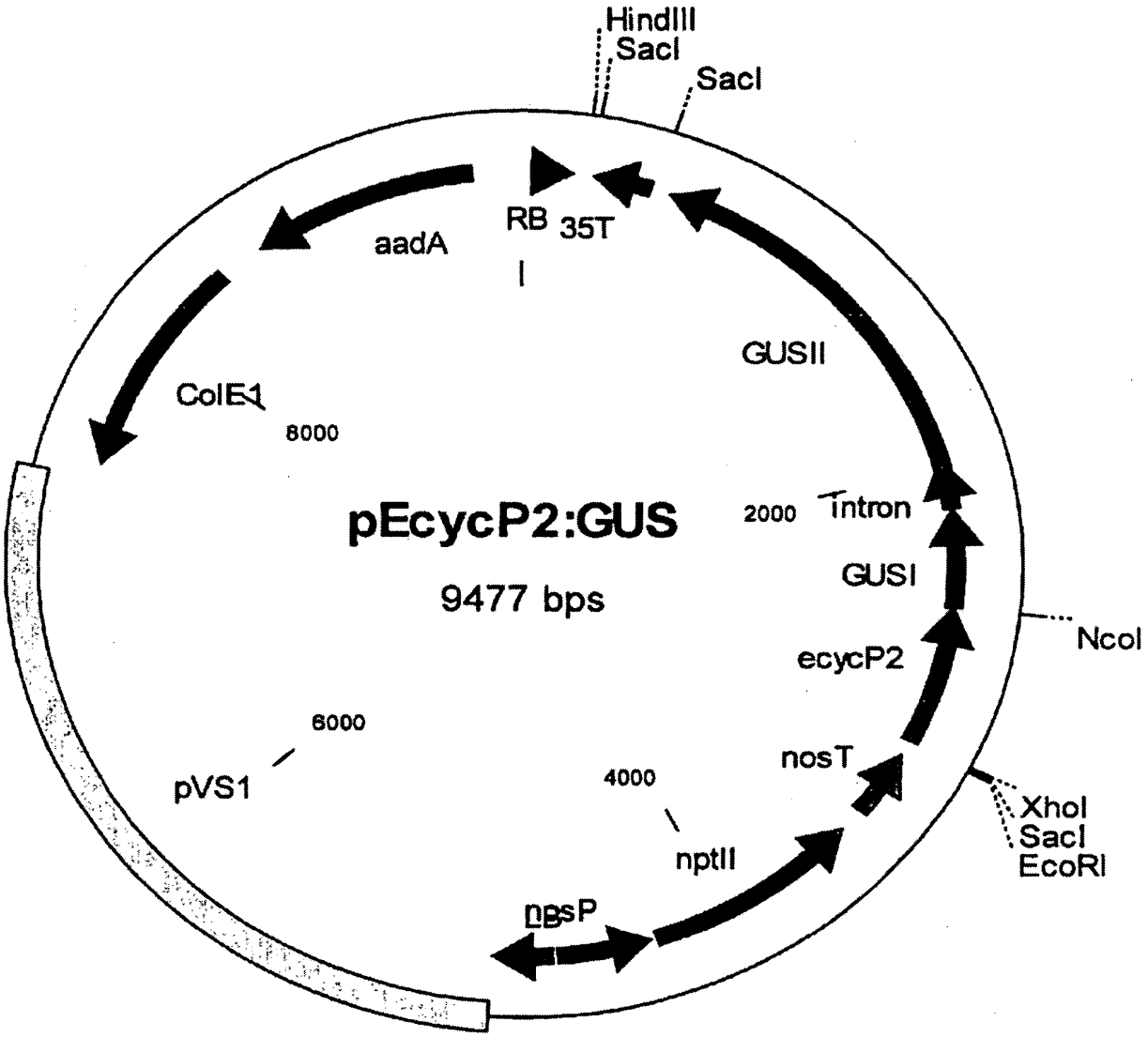

HindIII
SacI
SacI
RB
35T
aadA
GUSII
ColE1
8000
pEcycP2:GUS
9477 bps
2000
Intron
GUSI
NcoI
ecycP2
6000
nosT
pVS1
4000
XhoI
SacI
EcoRI
nptII
nosP

Fig.3

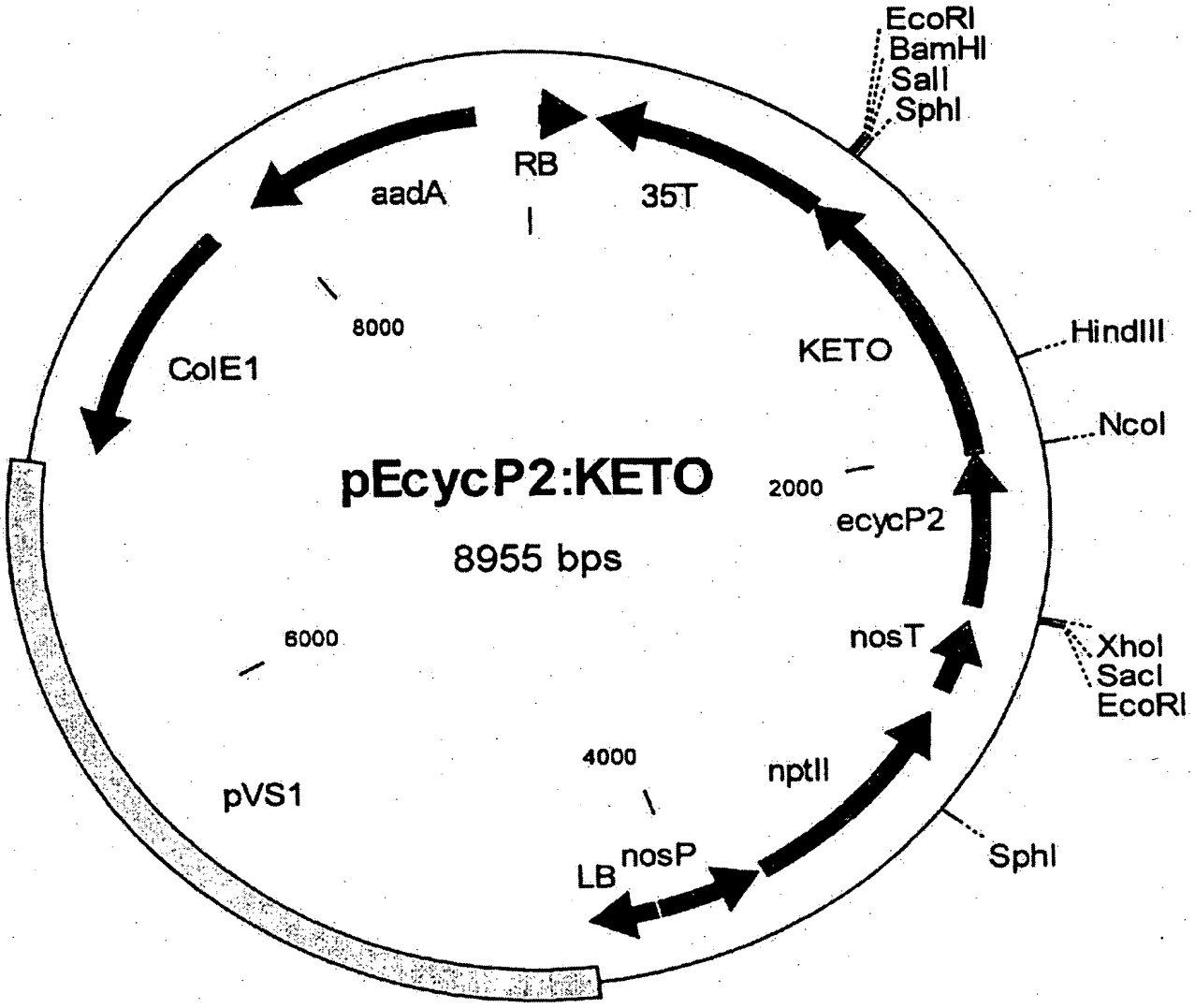

EcoRI
BamHI
SalI
SphI
RB
aadA
35T
8000
ColE1
KETO
HindIII
NcoI
pEcycP2:KETO
2000
ecycP2
8955 bps
6000
nosT
XhoI
SacI
EcoRI
4000
nptII
pVS1
LB
nosP
SphI

Fig.4

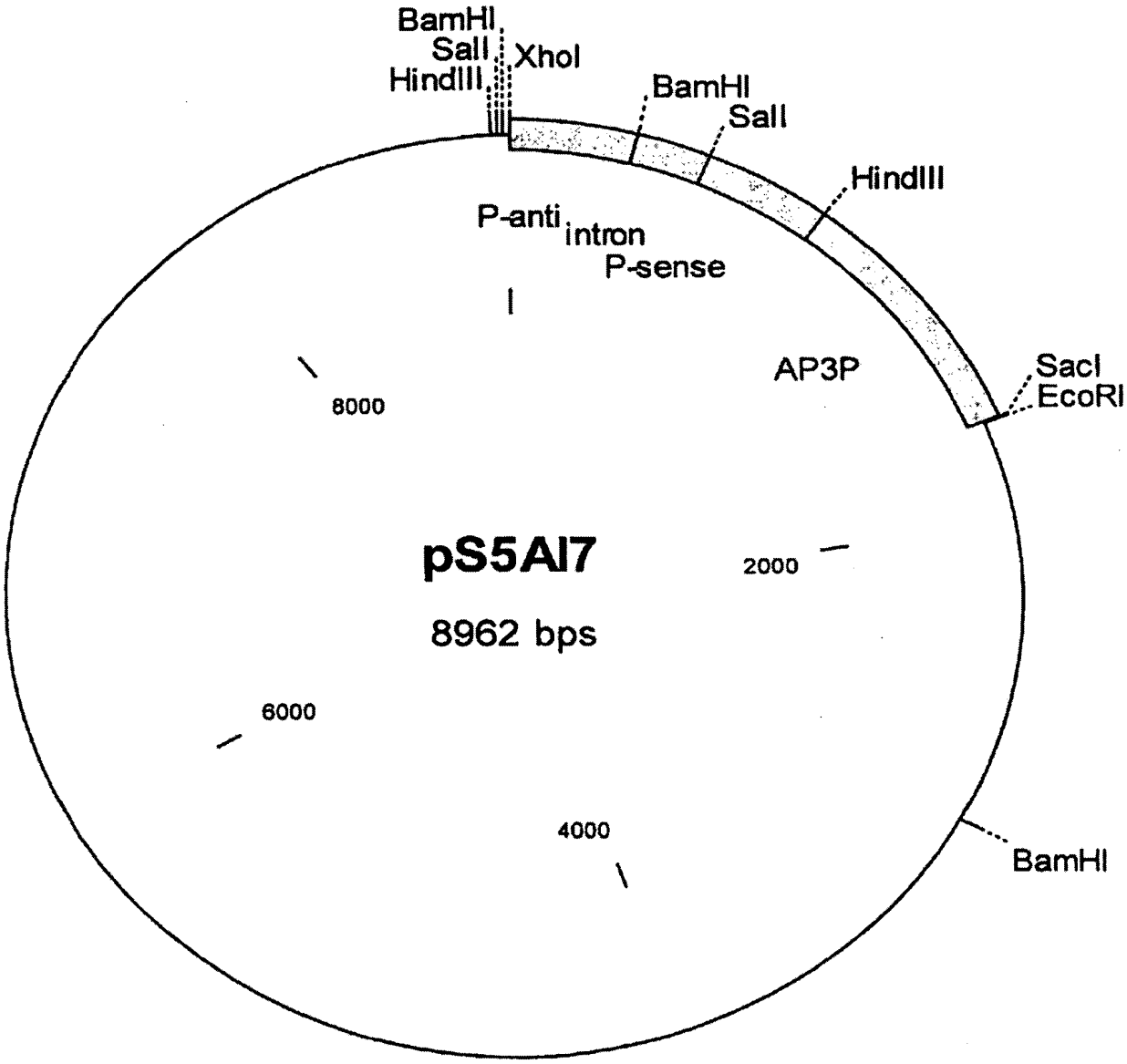

BamHI
SalI
HindIII
XhoI
BamHI
SalI
HindIII
P-anti
intron
P-sense
AP3P
SacI
EcoRI
8000
2000
pS5AI7
8962 bps
6000
4000
BamHI

Fig.5

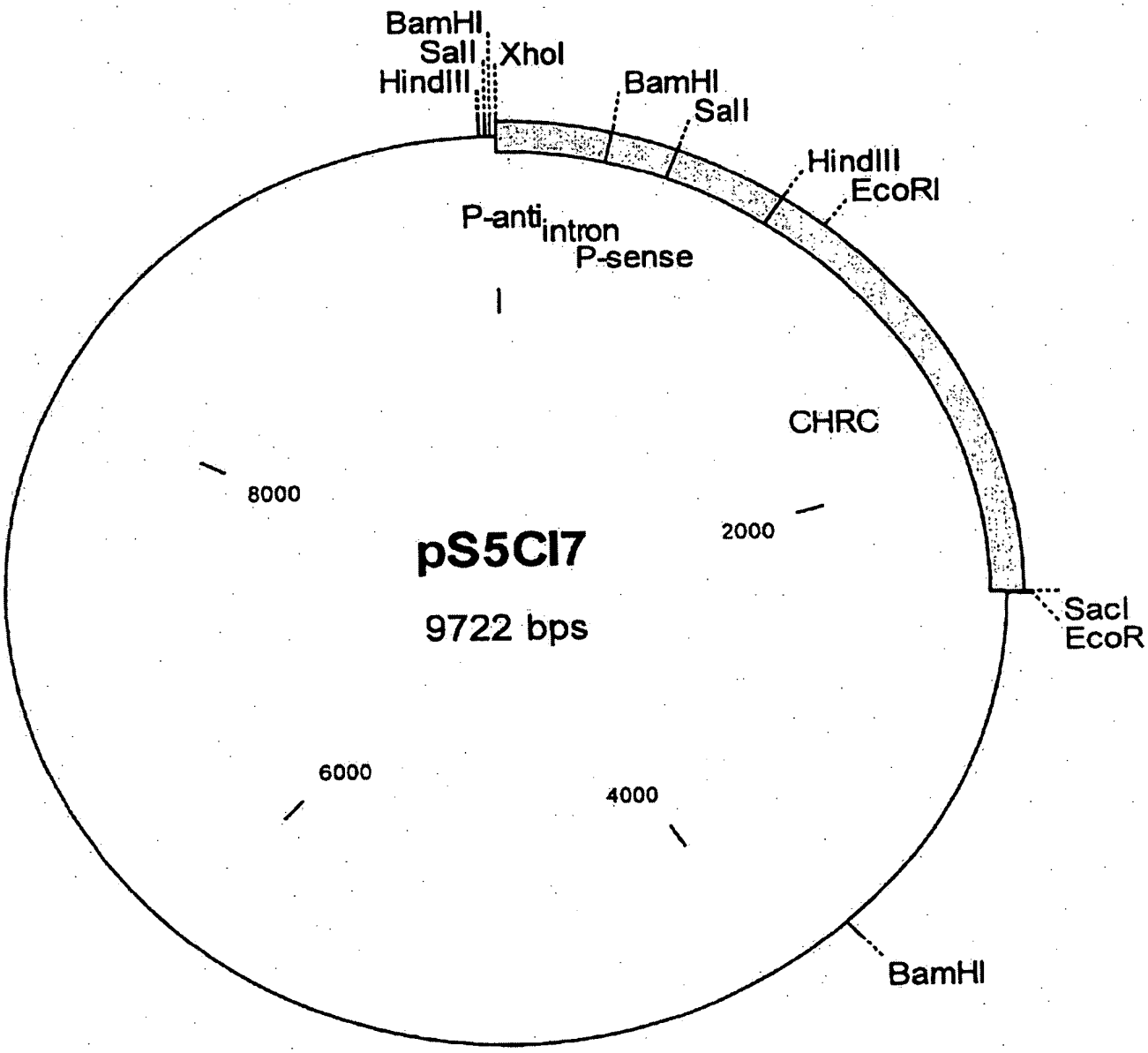
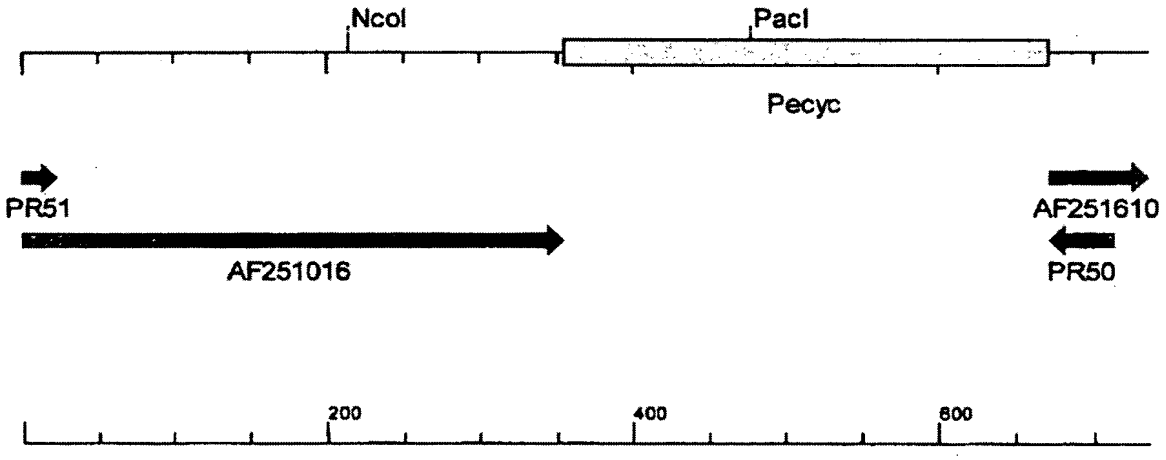
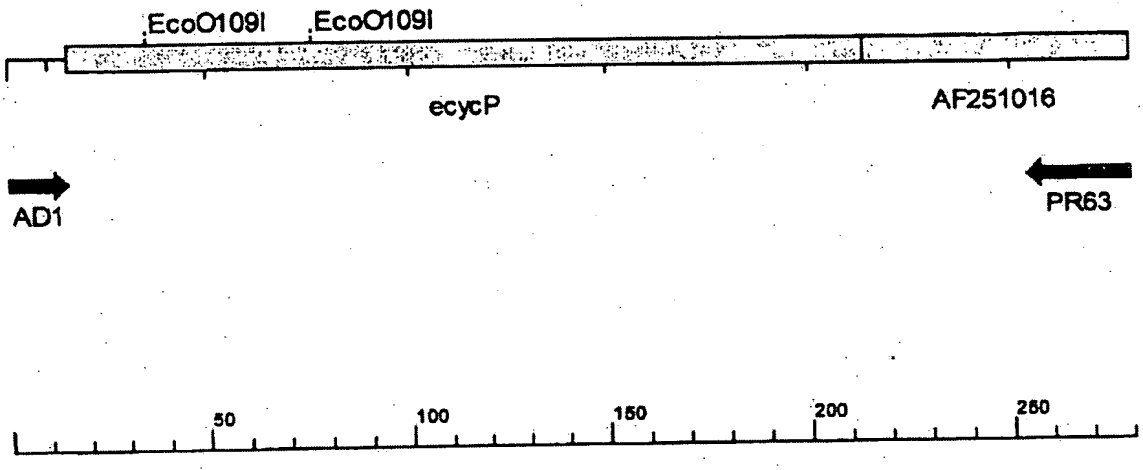

BamHI
SalI
XhoI
HindIII
BamHI
SalI
HindIII
EcoRI
P-anti
intron
P-sense
CHRC
8000
2000
pS5CI7
9722 bps
SacI
EcoR
6000
4000
BamHI

Fig.6



ecycP-IPCR (734 bps)

Fig.7



ecycP-TAIL (280 bps)

Fig.8

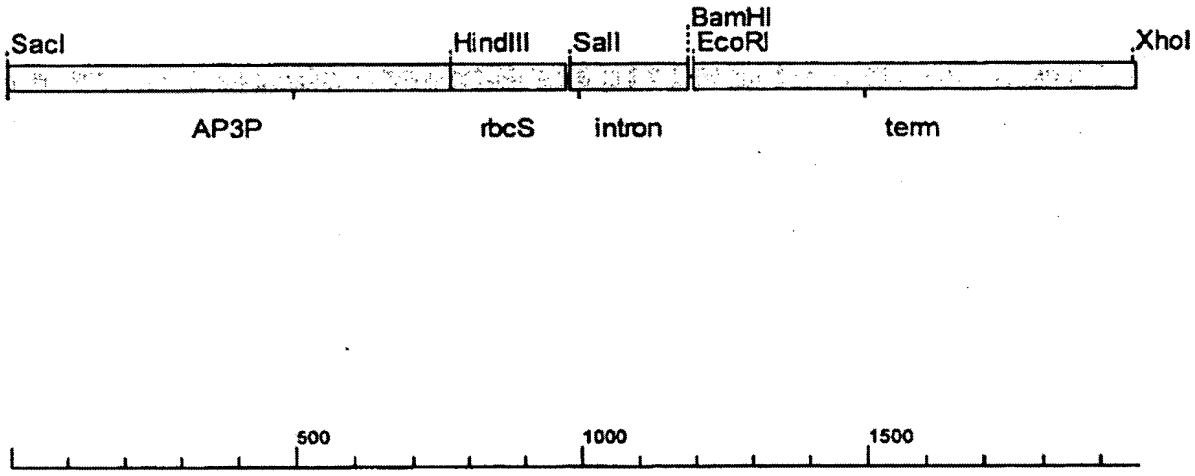
```
KETO2.seq  ATGCAGCTAGCAGCGACAGTAATGTTGGAGCAGCTTACCCGAAGCCCTGAGGCACTCAAGGAGAAGGAGAAGGAGGTTGCAGGCAGCTCTGACGTGTTGC 100
X86782.seq ATGCAGCTAGCAGCGACAGTAATGTTGGAGCAGCTTACCCGAAGCCCTGAGGCACTCAAGGAGAAGGAGAAGGAGGTTGCAGGCAGCTCTGACGTGTTGC 100

KETO2.seq  GTACATGGCCGACCCAGTACTCCCTTCCGTCAGAGGAGTCAGACCCGGCCCCCCCCCGGACTGAAGAATGCCTACAAGCCACCACCTTCCGACACAAAGGG 200
X86782.seq GTACATGGCCGACCCAGTACTCCCTTCCGTCAGAAGAGTCAGACCCGGCCCCCCCCCGGACTGAAGAATGCCTACAAGCCACCACCTTCCGACACAAAGGG 200

KETO2.seq  CATCACAATGGCCCTAGCTGTCATCGGCTCCTGGGCCCCAGTGTTCCTCCACGCCATTTTTCAAATCAAGCTTCCGACCTCCTTGGACCAGCTGCACTGG 300
X86782.seq CATCACAATGGCCCTACGTGTCATCGGCTCCTGGGCCCCAGTGTTCCTCCACGCCATTTTTCAAATCAAGCTTCCGACCTCCTTGGACCAGCTGCACTGG 300

KETO2.seq  CTGCCCGTGTCAGATGCCACAGCTCAGCTGGTTAGCGGCAGCAGCAGCCTGCTGCACATCGTCGTAGTATTCTTTGTCCTGGAGTTCCTGTACACAGGCC 400
X86782.seq CTGCCCGTGTCAGATGCCACAGCTCAGCTGGTTAGCGGCACGAGCAGCCTGCTGGACATCGTCGTAGTATTCTTTGTCCTGGAGTTCCTGTACACAGGCC 400

KETO2.seq  TTTTTATCACCACGCATGATGCTATGCATGGCACCATCGGCATGAGAAACAGGCAGCTTAATGACTTCTTGGGCAGAGTATGCATCTCCTTGTACGGCTG 500
X86782.seq TTTTTATCACCACGCATGATGCTATGCATGGCACCATCGGCATGAGAAACAGGCAGCTTAATGACTTCTTGGGCAGAGTATGCATCTCCTTGTACGGCTG 500

KETO2.seq  GTTTGATTACAACATGCTGCACGGCAAGCATTGGGAGCACCACAACCACACTGGCGAGGTGGGCAAGGACCCTGACTTCCACAGGGGAAACCCTGGCATT 600
X86782.seq GTTTGATTACAACATGCTGCACGGCAAGCATTGGGAGCACCACAACCACACTGGCGAGGTGGGCAAGGACCCTGACTTCCACAGGGGAAACCCTGGCATT 600

KETO2.seq  GTGGCCTGGTTTGCCAGCTTCATGTCCAGCTACATGTCGATGTGGCAGTTTGCCGCGCCTGGCATGGTGGACGGTGGTCATGCAGCTGCTGGGTGCGGCAA 700
X86782.seq GTGGCCTGGTTTGCCAGCTTCATGTCCAGCTACATGTCGATGTGGCAGTTTGCCGCGCCTGGCATGGTGGACGGTGGTCATGCAGCTGCTGGGTGCGGCAA 700

KETO2.seq  TGGCGAACCTGCTGGTGTTCATGGCGGCGCGGCCCATCCTGTCCGCCTTCCGCTTGTTCTACTTTGCCACGTACATGCCCCACAAGCCTGAGCCTGGCGC 800
X86782.seq TGGCGAACCTGCTGGTGTTCATGGCGGCGCGGCCCATCCTGTCCGCCTTCCGCTTGTTCTACTTTGCCACGTACATGCCCCACAAGCCTGAGCCTGGCGC 800

KETO2.seq  CCCGTCAGCCTCTTCACCAGCCGTCATGAACTGGTGGAAGTCCGCCACTAGCCAGCCGTCCGACCTGGTCAGCTTTCTGACCTGCTACCACTTCGACCTG 900
X86782.seq CCCGTCAGCCTCTTCACCAGCCGTCATGAACTGGTGGAAGTCCGCCACTAGCCAGCCGTCCGACCTGGTCAGCTTTCTGACCTGCTACCACTTCGACCTG 900

KETO2.seq  CACTCCGAGCACCACCCCTGGCCCTTTGCCCCCTGGTGGGAGCTGCCCAACTGCCCCCCGCCTGTCTGCCCGAGGTCTGGTTCCTGCCTAG          990
X86782.seq CACTCCGAGCACCACCCCTGGCCCTTCGCCCCCTGGTGGGAGCTGCCCAACTGCCCCCCGCCTGTCTGCCCGAGGTCTGGTTCCTGCCTAG          990
```


Fig.9

```
KETO2.pro  MQLAATVMLEQLTGSAEALKEKEKEVAGSSDVLRTWATQYSLPSEESDAA 50
X86782.pro MQLAATVMLEQLTGSAEALKEKEKEVAGSSDVLRTWATQYSLPSEESDAA 50

KETO2.pro  RPGLKNAYKPPPSDTKGITMALAVIGSWAAVFLHAIFQIKLPTSLDQLHW 100
X86782.pro RPGLKNAYKPPPSDTKGITMALRVIGSWAAVFLHAIFQIKLPTSLDQLHW 100

KETO2.pro  LPVSDATAQLVSGSSSLLHIVVVFFVLEFLYTGLFITTHDAMHGTIAMRN 150
X86782.pro LPVSDATAQLVSGTSSLLDIVVVFFVLEFLYTGLFITTHDAMHGTIAMRN 150

KETO2.pro  RQLNDFLGRVCISLYAWFDYNMLHRKHWEHHNHTGEVGKDPDFHRGNPGI 200
X86782.pro RQLNDFLGRVCISLYAWFDYNMLHRKHWEHHNHTGEVGKDPDFHRGNPGI 200

KETO2.pro  VPWFASFMSSYMSMWQFARLAWWTVVMQLLGAPMANLLVFMAAAPILSAF 250
X86782.pro VPWFASFMSSYMSMWQFARLAWWTVVMQLLGAPMANLLVFMAAAPILSAF 250

KETO2.pro  RLFYFGTYMPHKPEPGAASGSSPAVMNWWKSRTSQASDLVSFLTCYHFDL 300
X86782.pro RLFYFGTYMPHKPEPGAASGSSPAVMNWWKSRTSQASDLVSFLTCYHFDL 300

KETO2.pro  HWEHHRWPFAPWWELPNCRRLSGRGLVPA                      329
X86782.pro HWEHHRWPFAPWWELPNCRRLSGRGLVPA                      329
```

CA 02496300 2005-02-18

Fig.10



pJAI1 (1966 bps)

Fig.11

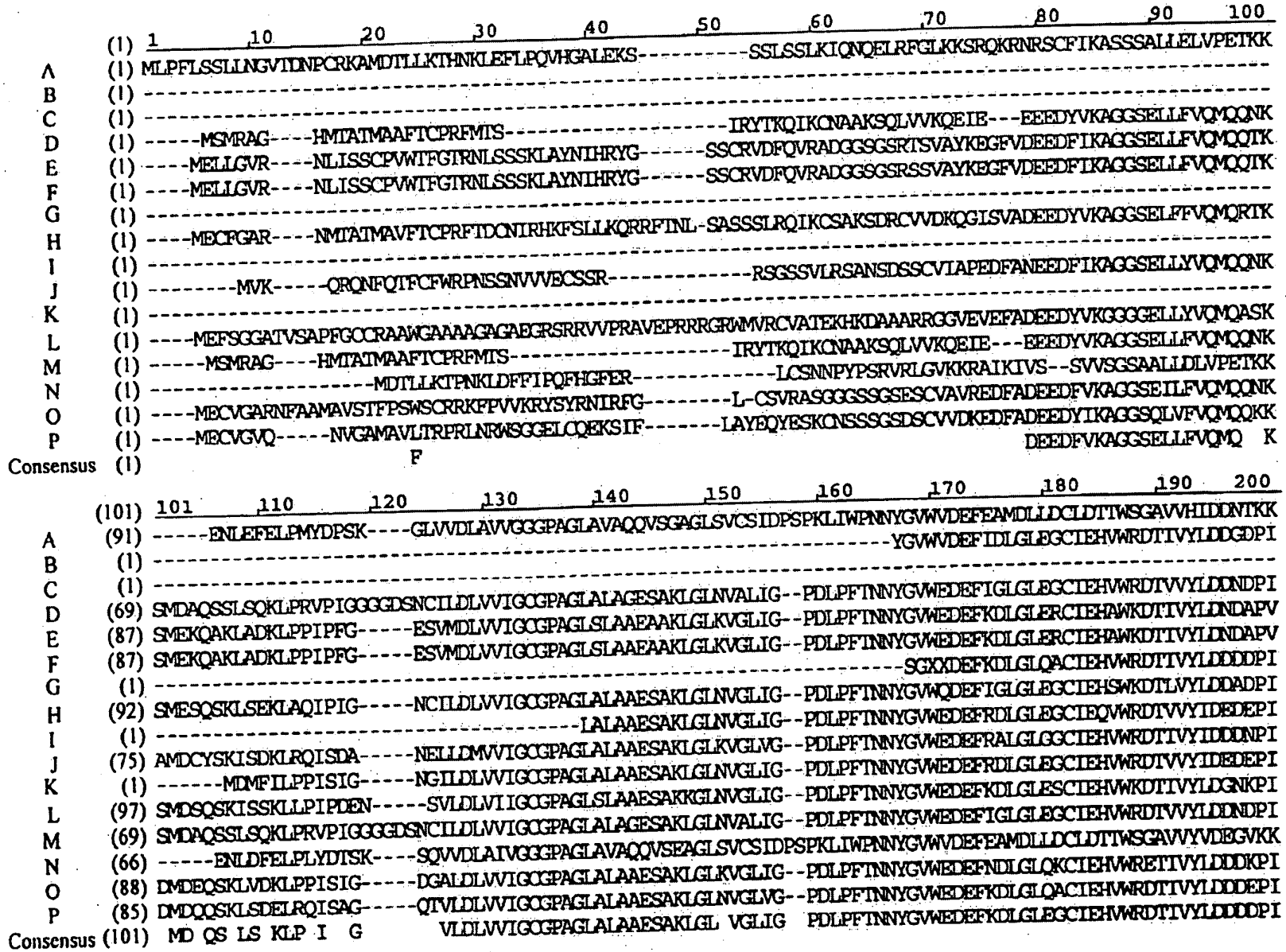


Fig.12A

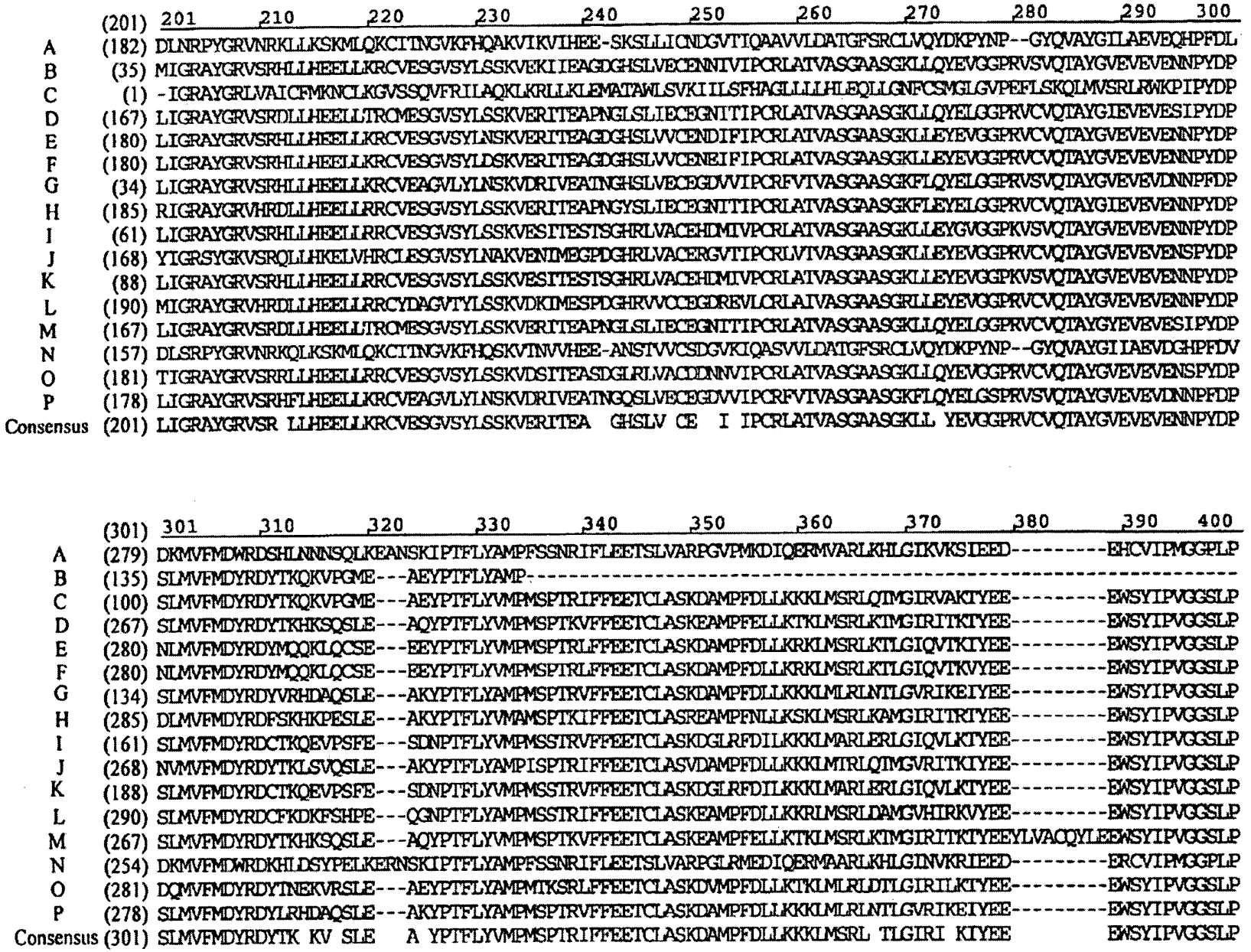


Fig.12B

```
                (401) 401       410       420       430       440       450       460       470       480       490       500
                      |         |         |         |         |         |         |         |         |         |         |
A               (370) VLPQRVVGIGGTAGMVHPSTGYMVARTLAAAPIVANAIVRSLSS---D-----RSLSGHKLSAEVWKDLWPIERRRQREFFCFGMDILLKLDLPATRRFF
B               (166) ----------------------------------------------------------------------------------------------------
C               (188) NTEQKNLAFGAAAR----------------------------------------------------------------------------------------
D               (355) NTEQKNLAFGAAASMVHPATGYSVVRSLSEAPNYAAVIAKILGKGNSKQMLDHGRYTTN-ISKQAWETLWPLERKRQRAFFLFGLALIVQMDIEGIRTFF
E               (368) NTEQKNLAFGAAASMVHPATGYSVVRSLSEAPKYASVIAKILKQDNSAYVVSGQSSAVN-ISMQAWSSLWPKERKRQRAFFLFGLELIVQLDIEATRTFF
F               (368) NTEQKNLAFGAAASMVHPATGYSVVRSLSEAPKYASVIAKILKQDNSAYVVSGQSSAVN-ISMQAWSSLWPKERKRQRAFFLFGLELIVQLDIEATRTFF
G               (222) NTEQKTLAFGAAASMVHPATGYSVVRSLSEAPKCAFVLANILRQNHSKNMLTSS-STPS-ISTQAWNTLWPQERKRQRSFFLFGLALILQLDIEGIRSFF
H               (373) NTEQKNLAFGAAASMVHPATGYSVVRSLSEAPNYAAVIAKILRQDQSKEMISLGKYTN--ISKQAWETLWPLERKRQRAFFLFGLSHIVLMDLEGIRTFF
I               (249) NTEQRNLAYGAAAS----------------------------------------------------------------------------------------
J               (356) NTEQRNLAFGAAASMVHPATGYSVVRSLSEAPKYASAIANLIKNDLSKNAILRQRSVGN-ISMQAWNTLWPQERKRQRAFFLFGLSLIVQLDIEGIRTFF
K               (276) NTEQRNLAFGAAASMVHPATGYSVVRSLSEAPNYASAIAYILKHDHSRGRLITHEQSNEN-ISMQAWNTLWPQERKRQRAFFLFGLALIIQLDIEGIRTFF
L               (378) NTDQKNLAFGAAASMVHPATGYSVVRSLSEAPRYASVISDILRNRVYPGEYLPGTSQSSSPSMLAWRTLWPQERKRQRSFFLFGLALIIQLNNEGIQTFF
M               (364) NTEQKNLAFGAAASMVHPATGYSVVRSLSEAPNYAAVIAKILGKGNSKQMLDLGRYTTN-ISKQAWETLWPLERKRQRAFFLFGLALIVQMDIEGIRTFF
N               (345) VLPQRVVGIGGTAGMVHPSTGYMVARTLAAAPIVANAIVRYLGSP-SS-----NSLRGDQLSAEVWRDLWPIERRRQREFFCFGMDILLKLDLDATRRFF
O               (369) NTEQKNLAFGAAASMVHPATGYSVVRSLSEAPKYASVIAEILREETT------KQINSN-ISRQAWDTLWPPERKRQRAFFLFGLALIVQFDIEGIRSFF
P               (366) NTEQKTLAFGAAASMVHPATGYSVVRSLSEAPKCASVLANILRQHYSKNMLTSS-SIPS-ISTQAWNTLWPQERKRQRSFFLFGLALILQLDIEGIRSFF
Consensus       (401) NTEQKNLAFGAAASMVHPATGYSVVRSLSEAP YASVIA ILR   S   L         IS QAW TLWP ERKRQRAFFLFGLALIVQLDIEG RTFF

                (501) 501       510       520       530       540       550       563
                      |         |         |         |         |         |         |
A               (462) DAFFDLEPRYWHGFLSSRLFLPELLVFGLSLFSHASNTSRLEIMAKGTLPLVNMINNLVQDTD
B               (166) ----------------------------------------------------------
C               (202) ----------------------------------------------------------
D               (454) RTFFRLPTWMWWGFLGSSLSSTDLIIFAFYMFIIAPHSLRMGLVRHLLSDPTGGIMLKAYLTI
E               (467) RTFFRLPTWMWWGFLGSSLSSFDLVLFSMYMFVLAPNSMRMSLVRHLLSDPSGAVMVKAYLER
F               (467) RTFFRLPTWMWWGFLGSSLSSFDLVLFSMYMFVLAPNSMRMSLVRHLLSDPSGAVMVRAYLER
G               (320) RAFFRVPKWMWQGFLGSSLSXADLMLFAFYMFIIAPNDMRRGLIRHLLSDPTGATLIRTYLTF
H               (471) RTFFRLPKWMWWGFLGSSLSSTDLIIFALYMFVIAPHSLRMELVRHLLSDPTGAIMVKAYLTI
I               (263) ----------------------------------------------------------
J               (455) RTFFRVPKWMWEGFLGSNLSSADLILFAFYMFFIAPNDLRMGLIRHLLSDPTGAIMIRTYITL
K               (375) RTFFRLPKWMWHGFLGSSLSSADLILFAFYMFIIAPNDLRKCLIRHLVSDPTGAIMVRTYLTL
L               (478) ETFFRLPKWMWRGFLGSTLSSVDLILFAFYMFTIAPNQMRMNLVRHLLSDPTGSIMIKTYLTL
M               (463) RTFFRLPTWMWWGFLGSSLSSTDLIIFAFYMFIIAPHSLRMGLVRHLLSDPTGGIMLKAYLTI
N               (439) DAFFDLQPHYWHGFLSSRLFLPELLVFGLSLFSHASNTSRLEIMTKGIVPLAKMINNLVQDRD
O               (462) RTFFRLPKWMWQGFLGSTLTSGDLVLFALYMFVISPNNLRKGLINHLISDPTGAIMIKTYLKV
P               (464) RAFFRVPKWMWQGFLGSSLSSADLMLFAFYMFIIAPNDMRKGLIRHLLSDPTGATLIRTYLTF
Consensus       (501) RTFFRLP WMW GFLGSSLSS DLILFA YMFIIAPN LRM LVRHLLSDPTGAIMIK YLT
```

Fig.12C

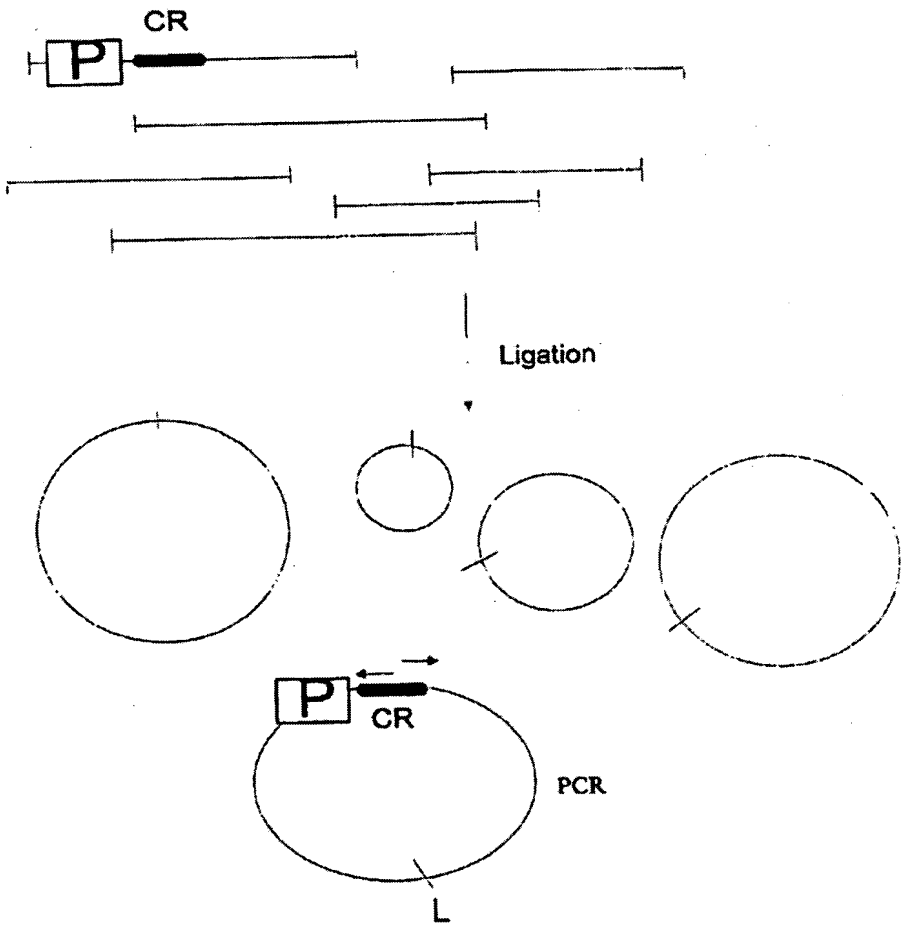


Fig. 13